

Ella Haarala

ORGANISAATION TOIMINTAA TUKEVAN DATANHALLINNAN KEHITTÄMINEN

Diplomityö
Johtamisen ja talouden tiedekunta
Tarkastaja: Maija Lampu
Tarkastaja: Pasi Hellsten
Maaliskuu 2025

TIIVISTELMÄ

Ella Haarala: Organisaation toimintaa tukevan datanhallinnan kehittäminen
Diplomityö
Tampereen yliopisto
Tietojohtamisen diplomi-insinöörin tutkinto-ohjelma
Maaliskuu 2025

Organisaatiot varastoivat dataa yhä suurempia määriä. Pelkkä datan olemassaolo ei kuitenkaan tuota organisaatioille arvoa, vaan dataa pitää pystyä hyödyntämään. Jotta organisaatiot voivat hyödyntää dataa ja luoda siitä arvoa, niiden täytyy kohdella sitä tärkeänä resurssina ja huolehtia sen laadusta sekä sääntelyn mukaisesta käsittelystä. Tässä tärkeässä osassa on datanhallinta. Datanhallinnalla viitataan dataomaisuuksien johtamiseen liittyvien päätöksentekotoimien ja vastuiden jakamiseen. Sen tarkoitus on maksimoida organisaation dataresurssien arvo. Datanhallinta kattaa lisäksi muun muassa dataomaisuuksien johtamiseen liittyvien käytäntöiden ja mittareiden määrittämisen, kommunikoinnin ja käyttöönoton. Datanhallinta ei ole kertaluontoisesti toteutettava projekti, vaan jatkuva prosessi, jonka täytyy mukautua sisäisiin ja ulkoihin muutoksiin. Datanhallinta tulisi aina muodostaa organisaatiokohtaisesti, sillä ei ole olemassa yhtä datanhallintamallia, jota voitaisiin hyödyntää tehokkaasti kaikissa organisaatioissa. Tämän haasteen on tunnistanut myös tutkimuksen kohdeyritys, kantaverkkoyhtiö Fingrid. Datanhallintamallia uudistaessaan Fingrid on todennut, että valmiit datanhallinnan viitekehitykset eivät usein vastaa organisaation tarpeita. Sen vuoksi organisaatio tarvitsisi uusia keinoja datanhallinnan kehittämiseen. Lisäksi kohdeorganisaatiossa on tunnistettu tarve selkeyttää nykytilaa kartoittamalla eri organisaatiotasojen näkemyksiä datanhallinnasta.

Tämä tutkimus pyrkii vastaamaan näihin haasteisiin selvittämällä, miten organisaatiot voivat kehittää datanhallintaa omaan toimintaansa sopivaksi. Tutkimuksessa kartoitetaan lisäksi kohdeorganisaation datanhallinnan nykytilaa, jotta voidaan arvioida, sopiiko tutkimuksessa muodostettu ratkaisu kohdeorganisaation datanhallinnan kehittämiseen. Tutkimus toteutetaan kirjallisuuskatsauksena ja empirisenä tapaustutkimuksena. Empiirisen tapaustutkimuksen aineistonkeruun menetelmänä toimii puolistrukturoidut haastattelut. Tutkimuksen tavoitteena on luoda malli, jonka avulla organisaatiot voivat kehittää datahallintamallin, joka tukee yrityksen toimintaa myös muutuvissa toimintaympäristöissä. Kehitysmalli muodostetaan sopivaksi organisaatioille, jotka haluavat uudistaa käytössä olevaa hallintamallia sekä organisaatioille, jotka ovat muodostamassa vasta ensimmäistä datanhallintamalliaan. Tutkimuksessa ei kuitenkaan keskitytä vain datanhallintamalliin, vaan datanhallintaa tarkastellaan kokonaisuutena, ottaen vahvasti huomioon myös sen toteuttaminen.

Tutkimukselle asetetut tavoitteet saavutettiin. Tutkimuksessa luotiin datanhallinnan jatkuva kehitysmalli, jonka rakenteen perustana käytettiin jatkuvan parantamisen Plan-Do-Check-Act-mallia. Mallin keskeisimmiksi osa-alueiksi määriteltiin organisaatiokohtaisen datanhallintamallin muodostaminen tai uudistaminen, datanhallinnan toteutus, datanhallinnan arviointi sekä toiminnan vakiinnuttaminen tai muutosten tekeminen. Organisaation toimintaa tukevan datanhallintamallin muodostamisessa tai uudistamisessa olennaisiksi asioiksi tunnistettiin liiketoiminstrategian sekä sisäisten ja ulkoisten kontekstitekijöiden yhteensopivuus datanhallintamallin kanssa. Datanhallinnan toteutuksen kannalta oleellisiksi edellytyksiksi tunnistettiin johdon tuki, rakenteiden ja periaatteiden jalkauttaminen, työntekijöiden sitouttaminen, henkilöstön kyvykkyydet, viestintä ja yhteistyö sekä aika- ja teknologiaresurssit. Lisäksi tutkimuksessa tunnistettiin kohdeorganisaation datanhallinnan vahvuuksia ja kehityskohteita. Näitä analysoimalla todettiin, että hyödyntämällä tutkimuksessa kehitettyä datanhallinnan kehitysmallia, kohdeorganisaatio voisi uudistaa datanhallintaansa tukemaan vielä paremmin organisaation toimintaa.

Avainsanat: datahallinta, datan operatiivinen hallinta, jatkuva kehitys, PDCA-malli

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin Originality Check -ohjelmalla.

ABSTRACT

Ella Haarala: Developing Data Governance to Support Organizational Operations
Master of Science Thesis
Tampere University
Master's Programme in Information and Knowledge Management
March 2025

Organizations are storing increasingly large amounts of data. However, the mere existence of data does not generate value for organizations; data must be utilized. To leverage data and create value from it, organizations must treat it as an important resource and ensure its quality and regulatory compliance. Data governance plays a crucial role in this. Data governance refers to the allocation of decision-making rights and responsibilities related to the management of data assets. Its purpose is to maximize the value of an organization's data resources. Data governance also includes defining, communicating, and implementing practices and metrics related to the management of data assets. Data governance is not a one-time project but a continuous process that must adapt to internal and external changes. Data governance should always be tailored to the organization, as there is no single data governance model that can be effectively utilized in all organizations. This challenge has also been recognized by the research target company, the transmission system operator Fingrid. In renewing its data governance model, Fingrid has found that ready-made data governance frameworks often do not meet the organization's needs, and therefore the organization needs new ways to develop data governance. Additionally, the target organization has identified the need to map the views of different organizational levels on data governance to clarify its current state.

This research aims to address these challenges by investigating how organizations can develop data governance to suit their business needs. The study also maps the current state of data governance in the target organization to assess whether the solution developed in the research is suitable for improving the target organization's data governance. The research is conducted as a literature review and an empirical case study, with semi-structured interviews as the data collection method. The goal of the research is to create a model that organizations can use to develop a data governance model that supports the company's business needs and operations even in changing environments. The development model is designed to be suitable for organizations that want to renew their existing governance model as well as for organizations that are forming their first data governance model. However, the research does not focus solely on the data governance model but considers data governance as a whole, strongly taking into account its implementation.

The research objectives were achieved. A continuous improvement model for data governance was created, based on the Plan-Do-Check-Act model of continuous improvement. The key areas of the model were defined as forming or renewing an organization-specific data governance model, implementing data governance, evaluating data governance, and stabilizing operations or making changes. In forming or renewing a data governance model that supports the organization's business and operations, the compatibility of the business strategy and internal and external contextual factors with the data governance model were identified as essential aspects. For the implementation of data governance, essential prerequisites were identified as management support, the deployment of structures and principles, employee engagement, staff capabilities, communication and collaboration, and time and technology resources. Additionally, the research identified the strengths and development areas of the target organization's data governance. By analyzing these, it was concluded that by utilizing the data governance development model created in the research, the target organization could renew its data governance to better support the organization's business and operations.

Keywords: data governance, data management, continuous improvement, PDCA-model

The originality of this thesis has been checked using the Turnitin Originality Check service.

TEKOÄLYN KÄYTTÖ OPINNÄYTTEESSÄ

Opinnäytteessäni on käytetty tekoälysovelluksia:

- Ei
- Kyllä

Ilmoitukseni mukaan olen käyttänyt opinnäytteessäni tutkielmanprosessin aikana seuraavia tekoälysovelluksia:

Scopus AI

ChatGPT GPT-4

Microsoft Copilot GPT-4

Käyttötarkoitus:

Scopus AI:ta hyödynnettiin kirjallisuuskatsauksen lähteiden etsimisen apuna. Ensiksi sovellukselle esitettiin työn aiheeseen liittyviä kysymyksiä. Sovellus listaa vastauksessaan käyttämänsä lähteet, joita tarkasteltiin ja hyödyllisiksi koetut lähteet otettiin kirjallisuuskatsaukseen mukaan. Sovellusta ei kuitenkaan hyödynnetty tekstin tuottamisessa.

ChatGPT:tä hyödynnettiin työssä kysely- ja haastattelurungon muodostamisen apuna. Kysymyksiä ei kuitenkaan otettu suoraan sovelluksen vastausten perusteella, vaan sitä käytettiin lähinnä ideoimisen apuna.

Copilot:ia käytettiin tukena englanninkielisten termien kääntämisessä. Copilotilla myös käännettiin työn tiivistelmä englanniksi.

Osiot, joissa tekoälyä on käytetty:

Abstract, luku 3 teoreettinen viitekehys ja luku 2.3.2 empiirisen aineiston keruu.

Olen tietoinen siitä, että olen täysin vastuussa koko opinnäytteeni sisällöstä, mukaan lukien osat, joissa on hyödynnetty tekoälyä, ja hyväksyn vastuun mahdollisista eettisten ohjeiden rikkomuksista.

ALKUSANAT

Diplomityön kirjoittaminen oli prosessina aikaa vievä ja välillä motivaatiota koetteleva, mutta samalla myös erittäin mielenkiintoinen ja palkitseva. Opin datanhallinnasta valtavasti ja työn kirjoittaminen toimi myös loistavana perehdytyksenä uusiin työtehtäviin. Harva diplomityö taitaa valmistua tavoiteaikataulussa ja myös itselläni palautus venyi työkiireiden vuoksi parilla kuukaudella. Olen prosessin aikana kuullut useammalta taholta, että valmis diplomityö on paras diplomityö. Olen ripustautunut tuohon ohjenuoraan vahvasti viimeisen kuukauden aikana, jotta työ tulisi valmiiksi eikä jäisi pöytälaatikkoon pölyttymään ja odottamaan viilauksia, joita ei koskaan tule. Ohje oli ilmeisen toimiva, sillä tässä se nyt sitten on: diplomityö parasta laatuaan, valmis sellainen.

Haluan kiittää työni ohjaajia Maija Lampua sekä Joonas Säynevirtaa hyvistä ehdotuksista, kommentteista ja tuesta kirjoitusprosessin aikana. Kiitos myös esihenkilölleni Hannu Hättöselle mahdollisuudesta tehdä Fingridille diplomityö mielenkiintoisesta aiheesta. Kiitos myös kaikille haastateltaville tutkimukseen osallistumisesta, ilman teitä ei diplomityötä olisi saatu aikaiseksi.

Lopuksi haluan kiittää perhettäni ja ystäviäni, jotka ovat tarjonneet tukea ja mukavia hetkiä opiskeluiden vastapainoksi näiden viiden vuoden ajan. Erityiskiitos Joonathanille, joka on kuunnellut opiskeluhöpinöitä ja -valitustani väsymättömästi sekä tuonut kaupasta motivaatiosuklaata huonoina päivinä.

Diplomityön valmistuminen päättää kahdeksan vuoden korkeakouluopintorupeamani, johon sisältyi tietojohdamisen opintojen lisäksi myös yksi ammattikorkeakoulututkinto. Vaikka opiskelija-alennuksista luopuminen kirpaiseekin, voin hyvillä mielin todeta, että nyt on tullut istuttua koulun penkillä tarpeeksi, ainakin toistaiseksi.

Järvenpäässä, 24.3.2025

Ella Haarala

SISÄLLYSLUETTELO

1. JOHDANTO	1
1.1 Tutkimuksen taustoitus ja aiempi tutkimus	1
1.2 Tutkimuksen tavoitteet ja tutkimuskysymykset	3
2. TUTKIMUSMETODOLOGIA	7
2.1 Tutkimusmetodologiset valinnat	7
2.2 Kirjallisuuskatsaus	10
2.3 Tapaustutkimus	15
2.3.1 Kohdeorganisaation esittely	15
2.3.2 Empiirisen aineiston keruu	17
2.3.3 Empiirisen aineiston analyysi	20
3. TEOREETTINEN VIITEKEHYS	23
3.1 Datanhallinta yleisesti	25
3.2 Organisaatiokohtaisen datanhallintamallin muodostaminen	26
3.2.1 Menettelylliset hallintamekanismit	28
3.2.2 Rakenteelliset hallintamekanismit	29
3.2.3 Päätöksentekoaalueet	34
3.3 Datanhallintamallin muodostamisessa huomioitavat taustatekijät	36
3.3.1 Datanhallinnan tavoitteet	37
3.3.2 Datanhallinnan muut kontekstitekijät	39
3.4 Datanhallinnan toteuttaminen	40
3.4.1 Johdon tuki	40
3.4.2 Työntekijöiden sitouttaminen datanhallintaan	41
3.4.3 Henkilöstön kyvykkyydet	42
3.4.4 Viestintä ja yhteistyö	42
3.5 Datanhallinnan arviointi	43
3.6 Kirjallisuuskatsauksen synteesi	44
4. EMPIIRISEN TUTKIMUKSEN TULOKSET	48
4.1 Taustoittavan kyselyn tulokset	48
4.2 Haastatteluiden tulokset	52
4.2.1 Liiketoimintalähtöisyys	53
4.2.2 Toimintamalleihin vaikuttavat tekijät	57
4.2.3 Datanhallinnan rakenteet	59
4.2.4 Datanhallinnan toteutus	63
4.2.5 Datanhallinnan hyödyt	69
5. JOHTOPÄÄTÖKSET	71
5.1 Jalostettu teoreettinen viitekehys	71
5.2 Kohdeorganisaation datanhallinnan vahvuudet ja kehityskohteet	77
6. YHTEENVETO	85
6.1 Tutkimuskysymyksiin vastaaminen	85
6.2 Tutkimuksen arviointi	86

6.3	Jatkotutkimusmahdollisuudet.....	88
	LÄHTEET.....	90
	LIITE A: KYSELYRUNKO	97
	LIITE B: HAASTATTELURUNKO	101

1. JOHDANTO

Tässä luvussa esitellään tutkimuksen tausta ja tavoitteet. Tutkimuksen tavoitteiden avulla muodostetaan tutkimuskysymykset. Lopuksi esitellään tutkimuksen rakenne.

1.1 Tutkimuksen taustoitus ja aiempi tutkimus

Teknologian kehitys on mahdollistanut organisaatioille yhä suurempien datamäärien keräämisen ja varastoinnin sekä monimuotoisemman datan hyödyntämisen. (Weibl & Hess 2019; Benfeldt et al. 2020; Bento et al. 2022) Samalla data on kehittynyt tietojärjestelmien sivutuotteesta yhdeksi organisaatioiden arvokkaimmista resursseista ja kilpailuedun lähteeksi (Tan et al. 2015; Weibl & Hess 2019; Ai et al. 2023). Pelkkä datan olemassaolo ei kuitenkaan riitä tuottamaan arvo ja kilpailuetua, vaan sitä pitää pystyä hyödyntämään (Otto 2011a). Data voi toimia kilpailuedun lähteenä niille organisaatioille, jotka pystyvät luomaan siitä arvoa esimerkiksi tietoperusteisen päätöksenteon, riskien hallinnan, kustannusten vähentämisen ja syvällisemmän liiketoiminnan ymmärtämisen kautta (Bento et al. 2022).

Jotta dataa voidaan hyödyntää arvon luomiseksi, organisaatioiden täytyy kohdella sitä arvokkaana resurssina ja huolehtia sen laadusta sekä sääntelyn mukaisesta käsittelystä (DeStefano et al. 2016). Datan laadun hallinta on tärkeää, koska hyvälaatuinen data luo pohjan liiketoiminnalle, dataan liittyvän sääntelyn noudattamiselle, tehokkaalle raportoinnille sekä integroiduille ja automatisoiduille liiketoimintaprosesseille (Otto 2011a). Dataan liittyvä sääntely taas on muodostunut organisaatioille yhä keskeisemmäksi muun muassa Euroopan Unionin yleisen tietosuoja-asetuksen, GDPR:n, voimaantulon jälkeen. GDPR:n noudattaminen vaatii, että organisaatiot tietävät mitä dataa on varastoitu, minne se on varastoitu ja miten tätä dataa käytetään. (Abraham et al. 2019) On todettu, että ilman selkeitä politiikkoja, sääntöjä ja kontroleja siitä, kuka vastaa mistäkin datasta, organisaation datan laatu heikkenee (Batini et al. 2009). Tämä taas vaikeuttaa säännösten noudattamista, koska esimerkiksi säännösten velvoittamaa raportointia ei välttämättä voida toteuttaa luotettavasti. Näiden tarpeiden myötä datanhallinnan merkitys on kasvanut organisaatioissa.

Datanhallinnalla (engl. data governance) viitataan dataomaisuuksien johtamiseen liittyvien päätöksenteko-oikeuksien ja vastuiden jakamiseen. Sen tarkoitus on maksimoida

organisaation dataresurssien arvo. (Otto 2011a) Datanhallinta kattaa lisäksi muun muassa dataomaisuuksien johtamiseen liittyvien käytänteiden ja mittareiden määrittämisen, kommunikoinnin ja käyttöönoton. Datanhallinta liittyy läheisesti datan operatiiviseen hallintaan (engl. data management). Datanhallinnan voidaan nähdä toteuttavan valvontaa ja varmistavan, että dataa hallitaan, kun taas datan operatiivinen hallinta toteuttaa tätä hallintaa määriteltyjen tavoitteiden saavuttamiseksi. (DAMA 2017) Datanhallinta ei ole kertaluontoisesti toteutettava projekti vaan jatkuva prosessi (DAMA 2017; Walsh et al. 2022), jota pitää pystyä sopeuttamaan ulkoisiin ja sisäisiin muutoksiin (DAMA 2017; Abraham et al. 2019; Mahanti 2019). Tämän vuoksi se vaatii myös jatkuvaa tarkastelua ja kehittämistä. Onnistunut datanhallinnan toteuttaminen mahdollistaa organisaatioille datan hyödyntämisen strategisena resurssina sekä pienentää riskejä liittyen datan väärinkäyttöön, heikkoon laatuun ja säännösten noudattamatta jättämiseen (Zhang 2022; Abraham et al. 2019). Lisäksi tutkimukset ovat osoittaneet yhteyden datan hallinnan ja yrityksen parantuneen kilpailukyvyn välillä (Abueed & Mehmet 2019; Medeiros et al. 2021). Vaikka datan hallinnan hyödyt tunnustetaan laajasti, on sen tehokas toteutus usein haastavaa (Benfeldt et al. 2020).

Tieteellisessä tutkimuksessa datan hallintaa on tutkittu paljon datan hallinnan määrittämisen kautta, mutta sen toteuttaminen käytännössä on jäänyt vähemmälle huomiolle (Benfeldt 2017; Alhassan et al. 2018; Al-Ruithe et al. 2019). Benfeldtin (2017) mukaan tarvittaisiin lisää tutkimusta siitä, miten datanhallinnan viitekehyksiä sovelletaan käytännössä ja missä kontekstissa datanhallintaa toteutetaan. Myös Abraham ja muut (2019) ovat artikkelissaan tunnistaneet tarpeen datanhallintamallien soveltamiselle. Heidän mukaansa lisää tutkimusta tarvittaisiin datanhallinnan vaikutuksista sekä kontekstitekijöistä, eli siitä, mitkä organisaatiokohtaiset tekijät vaikuttavat datanhallinnan suunnitteluun ja toteuttamiseen. Tämän näkökulman tärkeyttä tukee muissa tutkimuksissa tunnistettu ”one size fits all” lähestymistavan mahdottomuus. Tutkijat ovat todenneet, että ei ole olemassa yhtä datanhallintamallia, jota voitaisiin soveltaa kaikissa yrityksissä, vaan datanhallinta tulisi aina rakentaa organisaatiokohtaisesti (Weber et al. 2009; Al-Ruithe & Benkelifa 2017; Sothilingam et al. 2021). Sothilingam ja muut (2021) korostavat erityisesti datanhallinnan tavoitteiden merkitystä organisaatiokohtaisen datanhallinnan suunnittelussa. He väittävät, että suurin osa nykyisistä datanhallinnan lähestymistavoista ei auta organisaatioita määrittelemään, miten ne voivat saavuttaa datanhallinnan tavoitteet omassa liiketoimintaympäristössään. Organisaatiokohtaisen datanhallinnan kehittämisen tarve on tunnistettu myös tämän tutkimuksen kohdeorganisaatiossa, kantaverkko-yhtiö Fingridillä. Fingridillä on jo otettu käyttöön datanhallintamalli ja sen muodostamisessa ja uudistamisessa on todettu, että valmiit datanhallintamallit eivät usein vastaa

Fingridin tarpeita vaan ne ovat esimerkiksi liian raskaita tai painottuvat organisaation tavoitteiden kannalta vääriin asioihin.

Opinnäytetöissä datanhallintaa on aiemmin tutkittu muun muassa datan laadun kehittämisen (Hirvonen 2016; Sulanen 2021), datanhallinnan kypsyysmallien (Jokela 2023) sekä data-alustojen arvonluonnin (Tiilikainen 2023) kautta. Tämä tutkimus pyrkii täydentämään tätä tutkimuskenttää sekä vastaamaan edellä tunnistettuihin haasteisiin tarkastelemalla, mitä asioita organisaatioiden tulisi huomioida datanhallinnan muodostamisessa ja uudistamisessa, jotta datanhallinta saataisiin tukemaan organisaation toimintaa mahdollisimman hyvin. Tähän liittyy olennaisena osana käsitys datanhallinnasta jatkuvana prosessina, jota täytyy pystyä uudistamaan sisäisten ja ulkoisten muutosten mukana (DAMA 2017; Abraham et al. 2019; Mahanti 2019). Tutkimuksessa halutaan selvittää, onko datanhallinnan kehitykseen mahdollista löytää useammalle organisaatiolle sopivia suuntaviivoja. Tavoitteena ei siis ole yleistettävissä olevan datanhallintamallin kehitys. Sen sijaan tutkimuksessa pyritään luomaan datanhallinnan kehitysmalli, joka huomioisi organisaatiokohtaiset tarpeet ja ominaisuudet sekä niiden mahdollisen muuttumisen. Tarkoituksena on luoda kehitysmalli, jonka avulla organisaatiot voivat uudistaa tai luoda datahallintamallin, joka tukee yrityksen toimintaa myös muuttuvissa toimintaympäristöissä. Tutkimuksessa ei kuitenkaan keskitytä vain datanhallintamallin muodostamiseen, vaan datanhallintaa tarkastellaan kokonaisuutena, ottaen vahvasti huomioon myös sen toteuttaminen.

1.2 Tutkimuksen tavoitteet ja tutkimuskysymykset

Kuten edellä mainittu, tässä tutkimuksessa tarkastellaan datanhallintaa ja sen toteuttamista kohdeorganisaationa toimivan kantaverkkoyhtiö Fingridin kautta. Kohdeorganisaatiossa on erityisesti viime vuosina tunnistettu datanhallinnan tärkeys. Yhtiössä luotiin ensimmäiset datanhallinnan viralliset rakenteet vuonna 2019, minkä jälkeen organisaatiossa on tehty useita muita toimenpiteitä datanhallinnan kehittämiseksi. Vuonna 2024 kohdeorganisaatiossa tunnistettiin tarve uudelleen tarkastella ja määritellä datanhallinnan roolituksia. Tällöin nousi esiin jo aiemmin yhtiössä tunnistettu haaste valmiiden datanhallintamallien suorasta sopimattomuudesta yhtiön tilanteeseen ja datanhallinnan toimintaan. Yhtiössä todettiin, että useiden valmiiden mallien roolitukset ovat liian raskaita, eikä niitä haluta ottaa käyttöön sellaisenaan. Tämän seurauksena organisaatiossa syntyi halu ymmärtää paremmin, miten datanhallinnan viitekehyksiä voitaisiin soveltaa vastaamaan heidän tarpeisiinsa. Samalla kohdeorganisaatiossa tunnistettiin, ettei heillä ollut riittävän selkeää näkemystä siitä, kuinka hyödylliseksi tai organisaation tarpeisiin sopi-

vaksi nykyinen datanhallinta koetaan organisaatiossa yleisesti. Tämän vuoksi kohdeorganisaatiossa haluttiin kartoittaa eri organisaatiotasojen kokemuksia ja näkemyksiä datanhallinnasta nykytilanteen selkeyttämiseksi ja uusien kehityskohteiden kartoittamiseksi. Lisäksi haluttiin löytää keinoja, joilla datanhallinnan tilaa voitaisiin kehittää ja se saataisiin tukemaan yrityksen toimintaa mahdollisimman hyvin.

Tässä tutkimuksessa pyritään tuottamaan tietoa kohdeorganisaation datanhallinnan tilasta ja kehityskohteista, mutta myös datanhallinnan viitekehysistä, sekä tekijöistä, joihin organisaatioiden tulisi kiinnittää huomiota soveltaessaan näitä viitekehyskäytäntöjä ja muodostaessaan tai uudistaessaan toimintaansa sopivaa datanhallintaa. Tutkimus toteutetaan luomalla ensiksi katsaus datanhallintaa käsittelevään kirjallisuuteen, minkä perusteella rakennetaan työn teoreettinen viitekehys. Viitekehystä käytetään perustana empiiriseen tutkimukseen. Empiirisessä tutkimuksessa datanhallintaa tarkastellaan kohdeorganisaation kontekstissa tapaustutkimuksen avulla. Lopuksi empiirisen tutkimuksen tuloksia peilataan vielä teoreettiseen viitekehukseen.

Edellä esiteltyjen tekijöiden pohjalta työn päätutkimuskysymykseksi muodostettiin ”Miten organisaatio voi kehittää datanhallintaa omaan toimintaansa sopivaksi?”. Siihen pyritään vastaamaan seuraavilla alakysymyksillä:

1. Miten datanhallintamalli voidaan muodostaa tai uudistaa tietylle organisaatiolle sopivaksi?
2. Mitkä asiat tukevat datanhallinnan toteuttamista?
3. Mitkä ovat kohdeorganisaation nykyisen datanhallinnan vahvuudet ja heikoudet?

Kahteen ensimmäiseen alakysymykseen pyritään vastaamaan kirjallisuuskatsauksen avulla ja viimeiseen kysymykseen empiirisen tutkimuksen avulla. Kuvassa 1 on avattu tarkemmin, mitä tietoa työn teoreettisella ja empiirisellä osalla pyritään selvittämään.

Kirjallisuuskatsaus

- Mistä osa-alueista datanhallintamalli koostuu
- Mitkä tekijät tulee ottaa huomioon organisaatiokohtaisen datanhallintamallin suunnittelussa

Organisaatiokohtaisen datanhallintamallin muodostaminen / kehittäminen



- Mitkä asiat vaikuttavat siihen, onnistuuko datanhallinnan toteuttaminen
- Miten nämä tekijät tulisi ottaa huomioon datanhallinnan toteutuksessa

Datanhallinnan toteuttaminen



- Miksi datanhallinnan vaikutuksia tulisi seurata
- Miten vaikutuksia voidaan arvioida

Datanhallinnan arviointi

Empiirinen tutkimus

- Onko organisaation datanhallinnalla yhteys liiketoimintaan ja strategiaan
- Miten hyvin datanhallinta on omaksuttu organisaation yhteiseksi asiaksi ja pidetäänkö sitä tärkeänä
- Mitä ovat nykyisen mallin haasteet ja kehityskohteet

- Mikä on onnistunut kohdeorganisaation datanhallinnan toteutuksessa
- Mitä kehitettävää organisaatiolla on datanhallinnan toteutukseen liittyen

- Onko eri organisaatiotasolla erilaiset kokemukset datanhallinnan vaikutuksista
- Ovatko datanhallinnalla saavutettavat hyödyt selkeitä

Kuva 1. Työn teoreettisella ja empiirisellä osalla kartoitettava tieto

Työn lopputuloksena luodaan datanhallinnan jatkuvan parantamisen kehitysmalli, jonka avulla organisaatiot voivat kehittää organisaatiokohtaista, omaan toimintaansa sopivaa datanhallintaa. Kehitysmalli muodostetaan sopivaksi organisaatioiden ensimmäisen datanhallintamallin luomiseen sekä sen jatkuvaan kehittämiseen. Työssä selvitetään myös kohdeorganisaation datanhallinnan nykytilaa ja tunnistetaan siihen liittyviä kehityskohteita. Näiden avulla voidaan arvioida, sopiiko luotu malli kohdeorganisaation datanhallinnan kehittämiseen. Työn tekemiseen on varattu aikaa syyskuusta 2024 alkuvuoteen 2025 asti, joten keskeinen tutkimukseen vaikuttava rajoite on käytettävissä oleva aika. Tämän vuoksi työstä jätetään pois kehitysmallin laajempi testaaminen muissa yrityksissä. Tämä on kuitenkin olennainen jatkotutkimuskohde, sillä kehitysmalli luodaan osittain tapaustutkimuksen avulla, jolla ei lähtökohtaisesti pyritä laajoihin yleistyksiin. Tämän vuoksi on tärkeää testata mallia muissa konteksteissa ennen sen laajaa hyödyntämistä. Tutkimus alkaa johdannolla, jonka jälkeen toisessa luvussa kuvataan tutkimusmetodologiset valinnat sekä esitellään kirjallisuuskatsauksen ja tapaustutkimuksen toteutus. Kol-

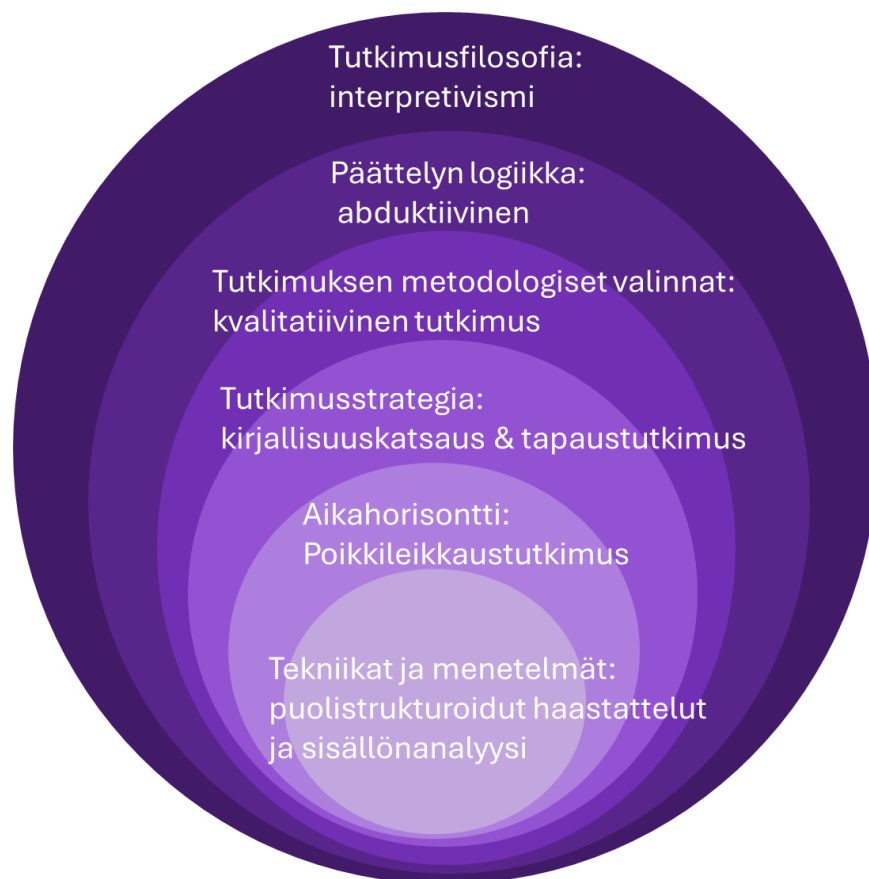
mannessa luvussa esitellään kirjallisuuskatsauksen tulokset ja luvun lopussa muodostetaan kirjallisuuskatsauksen synteesi. Tämän jälkeen neljännessä luvussa esitellään taustoittavan kyselyn sekä haastatteluiden tulokset. Viidennessä luvussa esitellään tutkimuksen johtopäätökset. Viimeisessä luvussa tehdään tutkimuksesta yhteenveto, arvioidaan tutkimusta sekä ehdotetaan jatkotutkimusmahdollisuuksia.

2. TUTKIMUSMETODOLOGIA

Tässä luvussa esitellään työssä tehdyt tutkimusmetodologiset valinnat. Tämän jälkeen kuvataan kirjallisuuskatsaus ja sen toteutus. Kirjallisuuskatsauksen kuvaamisen jälkeen tutustutaan tapaustutkimukseen ja sen toteutukseen.

2.1 Tutkimusmetodologiset valinnat

Tutkimusprosessia voidaan jäsentää ja selkeyttää niin kutsutun Saundersin tutkimussipulin avulla, jossa erilaisia tutkimuksen elementtejä kuvataan kuutena kerroksena. Näitä kerroksia ovat tieteenfilosofia, päättelyn logiikka, tutkimuksen metodologiset valinnat, tutkimusstrategia, tutkimuksen aikahorisontti sekä tekniikat ja menetelmät. (Saunders et al. 2019) Kuvassa 2 on esitetty tässä tutkimuksessa tehdyt tutkimusmetodologiset valinnat.



Kuva 2. Tutkimusmetodologian kuvaus Saundersin tutkimussipulin avulla (mukailen Saunders et al. 2019)

Tieteenfilosofialla tarkoitetaan tutkimuksen taustalla olevia uskomuksia ja oletuksia tiedon kehittämisestä. Ne sisältävät ontologisia olettamuksia todellisuuden luonteesta,

epistemologisia oletuksia tiedon luonteesta sekä oletuksia tutkijan arvojen vaikutuksesta tutkimusprosessiin. (Saunders et al. 2019) Tässä tutkimuksessa tieteenfilosofiana toimii interpretivismi. Saundersin ja muiden (2019) mukaan interpretivismisen tutkimuksen tarkoitus on luoda uutta, rikkaampaa ymmärrystä ja tulkintoja sosiaalisesta maailmasta ja konteksteista. He esittävät, että liiketoiminnan ja johtamisen tutkimuksessa tämä tarkoittaa organisaation tarkastelua eri ihmisryhmien näkökulmasta. Tässä työssä eri ihmisryhmien näkökulmien tarkastelu koetaan tärkeäksi, koska tutkimuksessa halutaan selvittää eri organisaatiotasojen näkemyksiä ja kokemuksia datanhallinnasta ja käyttää niitä muun muassa haasteiden ja kehityskohtien tunnistamiseen. Interpretivismiselle tutkimukselle on myös ominaista tutkijan tekemien tulkintojen merkitys, pienet otokset, syvälinen analyysi sekä laadulliset menetelmät (Saunders et al. 2019). Ne ovat olennaisia tässä tutkimuksessa.

Tutkimuksen päättelyn logiikka eli lähestymistapa teorian kehittämiseen voi olla joko deduktiivinen, induktiivinen tai abduktiivinen. Deduktiivinen päättely lähtee liikkeelle teoriasta. Sen tavoitteena on vahvistaa tai kumota hypoteeseja ja teorioita. Induktiivinen päättely taas alkaa havainnoista ja tutkimusaineistosta ja sen tavoitteena on luoda ja rakentaa uusia teorioita. Abduktiivinen päättely on näiden kahden lähestymistavan yhdistelmä. Sen tavoitteena on uuden teorian luominen tai muokkaaminen. (Saunders et al. 2019) Tässä tutkimuksessa noudatetaan abduktiivista päättelyä. Se koettiin tutkimuksen kannalta parhaaksi lähestymistavaksi, koska sen lähtökohtiin kuuluu olemassa olevan teorian kehittäminen. Tutkimuksessa luodaan kirjallisuuden avulla teoreettinen viitekehys, jota käytetään empiirisen tutkimuksen rakentamiseen ja analysointiin. Eli työ lähtee teoriasta, jonka avulla luodaan empiirinen osuus. Empiirisen tutkimuksen toteuttamisen jälkeen sen tuloksia peilataan vielä takaisin teoriaan, jotta teoreettista viitekehystä voidaan täydentää sekä tunnistaa kohdeorganisaation datanhallinnasta olennaisia kehityskohteita ja esittää kehitysehdotuksia. Työn tarkoituksena on siis kehittää ja soveltaa olemassa olevaa teoriaa yhden yrityksen syvällisellä analyysillä, ei testata teorian paikansäilyvyyttä tai luoda täysin uutta teoriaa.

Tutkimuksen metodologiset valinnat kuvaavat, käytetäänkö tutkimuksessa kvalitatiivisia menetelmiä, kvantitatiivisia menetelmiä vai niiden yhdistelmää. Tämä tutkimus on kvalitatiivinen eli laadullinen tutkimus. Saundersin ja muiden (2019) mukaan kvalitatiivisen ja kvantitatiivisen tutkimuksen erottaa toisistaan ei-numeerinen ja numeerinen data. Tutkijoiden mukaan laadullista tutkimusta käytetään usein synonyymina datan keräys- ja analyysimenetelmille, jotka tuottavat tai käyttävät ei-numeerista dataa. Eskolan ja Suorannan (2008) mukaan laadulliselle tutkimukselle on tyypillistä harkinnanvarainen otanta eli tutkimuksessa keskitytään usein pieneen määrään tapauksia ja pyritään analysoimaan

niitä perusteellisesti. Aineiston tieteellisyyden kriteeri ei ole aineiston määrä vaan sen laatu.

Tutkimusstrategia on suunnitelma, joka kuvaa, miten tutkija aikoo vastata tutkimuskysymykseensä (Saunders et al. 2019). Tämän laadullisen tutkimuksen tutkimusstrategiana toimii kirjallisuuskatsaus ja tapaustutkimus. Finkin (2005) mukaan kirjallisuuskatsaus on täsmällinen, systemaattinen ja toistettavissa oleva menetelmä, jolla tunnistetaan, arvioidaan ja syntetisoidaan aikaisempaa tutkimusta. Kirjallisuuskatsaus päätettiin yhdistää tapaustutkimukseen, sillä kirjallisuuskatsauksen avulla luodaan työn teoreettinen viitekehys, jota käytetään tapaustutkimuksen perustana ja siihen peilataan tapaustutkimuksen tuloksia. Tapaustutkimuksessa teoreettisen viitekehksen rakentaminen on olennaista. Voss ja muut (2002) argumentoivat, että ilman teoriaa empiirisesti kerätystä datasta on mahdotonta luoda ymmärrystä, koska positiivisia ja negatiivisia tuloksia ei voida erottaa toisistaan. Kirjallisuuskatsauksen prosessi on kuvattu tarkemmin alaluvussa 2.2.

Saundersin ja muiden (2019) mukaan tapaustutkimus on syvälinen tutkimus ilmiöön sen todellisessa ympäristössä. Tapaustutkimuksessa tutkimusasetelma muodostetaan yhden tutkittavaa ilmiötä edustavan tapauksen tai valikoitujen tapausten varaan (Vuori 2021b). Tässä tutkimuksessa tapauksena toimii kohdeorganisaation datanhallinta. Vuoren (2021b) mukaan tapaustutkimuksessa tuotetaan tutkimuskohteesta tarkka ja havainnollinen kuvaus, jonka avulla pyritään oppimaan uutta tarkasteltavasta ilmiöstä ja soveltamaan tietoa myös muissa yhteyksissä. Tapaustutkimuksessa ei kuitenkaan pyritä laajoihin yleistyksiin. Tapaustutkimus valittiin toiseksi tutkimusstrategiaksi, koska tutkimuksessa on tarve tarkastella datanhallintaa syvällisesti nimenomaan kohdeorganisaation kontekstissa. Johdannossa todettiin, että ei ole olemassa yhtä datanhallintamallia, jota voitaisiin soveltaa kaikissa yrityksissä vaan datanhallinta tulisi aina rakentaa organisaatiokohtaisesti (Weber et al. 2009; Al-Ruithe & Benkhelifa 2017; Sothilingam et al. 2021). Tämän perusteella ilmiötä on perusteltua tutkia yhden organisaation kautta tapaustutkimuksen avulla. Tapaustutkimus on myös hyvä lähestymistapa tutkimukseen, jossa pyritään vastaamaan ”miten” ja ”miksi” kysymyksiin (Baxter & Jack 2008; Farquhar 2012). Myös työn tutkimuskysymyksen asettelu ”Miten organisaatio voi kehittää datanhallintaa omaan toimintaansa sopivaksi?” tukee siis tapaustutkimuksen valintaa työn tutkimusstrategiaksi. Tapaustutkimusta avataan lisää alaluvussa 2.3.

Aikahorisontti kuvaa ajanjaksoa, jonka aikana tutkimus toteutetaan. Tässä tutkimuksessa aikahorisonttia kuvaa poikkileikkaustutkimus. Poikkileikkaustutkimukset sijoittuvat tiettyyn rajalliseen ajanhetkeen. Poikkileikkaustutkimuksessa tarkastellaan ilmiötä tietyn ajanhetkenä, kun taas pitkittäistutkimuksessa ilmiön kehittymistä tarkastellaan pidemmän ajan kuluessa. Poikkileikkaustutkimuksessa poikkileikkauksia voi myös olla

useita. (Eriksson & Koistinen 2005) Tässä tutkimuksessa tarkastellaan yrityksen datanhallinnan tilaa loppuvuodesta 2024 eli poikkileikkauksia on vain yksi.

Valitut tekniikat ja menetelmät kuvaavat työssä käytettyjä aineistonkeruu ja -analyysimenetelmiä. (Saunders et al. 2019). Tässä tutkimuksessa aineistonkeruumenetelmänä käytetään puolistrukturoituja haastatteluita. Haastatteluihin yhdistetään lisäksi taustoittava kysely, joka lähetetään haastateltaville ennen haastattelua. Taustoittava kysely päätettiin yhdistää haastatteluihin, sillä organisaatiossa haluttiin saada helposti vertailtavaa ja esitettävää tietoa haastatteluiden lisäksi. Aineiston analyysimenetelmänä käytetään sisällysanalyysia. Tutkimuksessa käytetyt empiiriseen tutkimukseen liittyvät menetelmät kuvataan tarkemmin luvussa 2.3.

2.2 Kirjallisuuskatsaus

Tässä tutkimuksessa kirjallisuuskatsaus toteutettiin Finkin (2005) ehdottaman seitsemänvaiheisen prosessimallin avulla. Prosessimalliin kuuluvat seuraavat vaiheet:

1. Tutkimuskysymysten valinta
2. Tietokantojen ja lähteiden valinta
3. Hakutermien valinta
4. Käytännön seulan asettaminen
5. Metodologinen rajaus
6. Katsauksen tekeminen
7. Tulosten syntetisointi

Työn tutkimuskysymykset asetettiin jo edellisessä luvussa. Ennen tutkimuskysymysten asettamista käytiin keskustelua kohdeorganisaation kanssa ja tutkittiin alustavasti aiheeseen liittyvää kirjallisuutta. Varsinainen tietokantojen ja lähteiden valinta tehtiin kuitenkin vasta tutkimuskysymysten määrittämisen jälkeen. Tehdyt valinnat sekä prosessin kohdat 3, 4 ja 5 on kuvattu seuraavissa kappaleissa. Prosessin viimeinen vaihe ”Tulosten syntetisointi” eli löydösten selittäminen ja tutkimuksen luotettavuuden arviointi toteutettiin katsauksen jälkeen.

Kirjallisuuskatsauksessa hyödynnettäviksi tietokannoiksi valittiin Scopus, Google Scholar, sekä Tampereen Yliopiston tietokanta Andor. Lisäksi aineiston etsimisessä käytettiin luonnollisen kielen prosessointiin perustuvaa Scopus-tietokannan lisäosaa, joka analysoi tietokannan julkaisuja ja luo vastauksen niiden tiivistelmien perusteella. Työkalu lis-

taa myös käyttämänsä lähteet. Scopus AI -lisäosaa hyödynnettiin kirjallisuuden etsimiseen tarkastelemalla vastauksissa käytettyjä lähteitä. Työkalua ei käytetty itse tekstin tuottamiseen.

Katsauksessa käytetyt hakutermit ja Scopus AI:lle annetut kehotukset ja kysymykset eli promptit muodostettiin työn keskeisempien käsitteiden pohjalta. Keskeiset käsitteet on lueteltu alla suomeksi ja englanniksi, koska haku toteutettiin pääasiassa englanniksi.

- data
- datanhallinta (engl. data governance)
- datan operatiivinen hallinta (engl. data management)
- PDCA-malli (engl. PDCA-cycle)
- päätöksentekoaalueet (engl. decision domains)
- roolit ja vastuut (engl. roles and responsibilities)
- datanhallinnan tavoitteet (engl. data governance objectives)
- kontekstitekijät (engl. contingency factors)
- datanhallinnan toteuttaminen (engl. implementation of data governance)
- vaikutukset (engl. effects)

Näiden käsitteiden avulla luodut Scopus AI:lle annetut keskeisimmät kehotteet on myös listattu alla. Kehotteet annettiin englanniksi, koska todettiin, että työkalu toimii parhaiten englannin kielellä.

- Data as a source of competitive advantage and value creation
- What types of data governance frameworks exist?
- Data governance and PDCA
- What are the objectives of data governance
- How are contingency factors related to data governance
- What is the relationship between data governance and business strategy and objectives?
- Why is the implementation of data governance difficult

Keskeisimmät hakutermit ja niillä saadut rajoittamattoman haun tulokset on esitetty taulukossa 1. Hakutermit on esitetty siinä muodossa, missä niitä on käytetty Andor-tietokannassa. Esimerkiksi Google Scholar ei tue kaikkien samojen erikoismerkkien käyttöä, joten hakutermin muotoilu on sovitettu kullekin tietokannalle sopivaksi.

Taulukko 1. *Rajoittamattoman haun tulokset*

Hakutermi	Andor	Google Scholar	Scopus	Yhteensä
"data governance" AND definition*	396	44 300	103	44 799
"data governance" AND (model OR framework)	3 877	1 605	56 700	62 182
"data governance" AND ("decision domains" OR "decision areas" OR "responsibility domains")	15	420	15	450
"data governance" AND (goals OR objective*)	1 175	38 900	516	40 591
"data governance" AND (antecedents OR contingency)	46	5 940	28	6 014
"data governance" AND implement*	2 277	44 400	763	47 440
"data governance" AND (outcomes OR effect* OR impact)	3 070	44 600	1 147	48 817
"data governance" AND (challenge* OR barrier* OR obstacle* OR difficult* OR problem*) AND (implement* OR adoption OR rollout OR integration)	1 355	36 100	571	38 026
"data governance" AND success	630	37 600	132	38 362
"data governance" AND (measur* OR metric* OR evaluat*)	2 220	39 100	670	41 990

Kuten taulukosta 1 nähdään, rajoittamaton haku tuotti pääasiassa erittäin suuren määrän tuloksia. Tämän takia kirjallisuuskatsauksessa toteutettiin Finkin prosessimallin neljäs vaihe, käytännön seulan asettaminen. Haussa rajoitettiin tulokset vuoden 2015 jälkeen julkaistuun aineistoon. Lisäksi aineisto rajattiin englannin- ja suomenkieliseen aineistoon Andorin ja Scopusin osalta. Google Scholar ei tukenut vuosirajoituksen lisäksi muiden relevanttien rajoitusten asettamista. Scopusissa aineisto rajattiin vielä artikkeleihin ja konferenssijulkaisuihin, kun taas Andorissa aineisto saatiin rajattua vertaisarvioituihin lehtiin. Taulukossa 2. on kuvattu rajoitetun haun tulokset.

Taulukko 2. *Rajoitetun haun tulokset*

Hakutermi	Andor	Google Scholar	Scopus	Yhteensä
"data governance" AND definition*	63	17 500	77	17 640
"data governance" AND (model OR framework)	979	29 600	1 256	31 835
"data governance" AND ("decision domains" OR "decision areas" OR "responsibility domains")	9	341	10	360
"data governance" AND (goals OR objective*)	399	20 400	378	21 177
"data governance" AND (antecedents OR contingency)	19	5 370	22	5 411

"data governance" AND implement*	485	17 700	565	18 750
"data governance" AND (outcomes OR effect* OR impact)	799	22 300	820	23 919
"data governance" AND (challenge* OR barrier* OR obstacle* OR difficult* OR problem*) AND (implement* OR adoption OR rollout OR integration)	402	17 000	431	17 833
"data governance" AND success	86	17 500	86	17 672
"data governance" AND (measur* OR metric* OR evaluat*)	460	20 500	481	21 441

Taulukosta 2. nähdään, että rajoitetun haun tulokset olivat vielä osin niin suuria, ettei niitä kaikkia ollut mahdollista käydä läpi tutkimuksen aikarajoituksen puitteissa. Tämän vuoksi aineistoa saatettiin vielä rajata Scopuksessa tietyn avainsanan mukaan tai Andorissa tietyn aihepiirin mukaan. Näin saatiin rajattua pois esimerkiksi liian IT-painotteisia julkaisuja. Avainsanoina käytettiin esimerkiksi "data governance", "data quality" ja "decision making". Lisäksi haulla saadusta aineistosta päätettiin tarkastella 50:tä ensimmäistä hakutulosta. Näitä hakutuloksia arvioitiin ensiksi otsikkotasolla. Otsikon perusteella arvioitiin, onko aihe niin relevantti tutkimuksen kannalta, että aineiston tiivistelmä kannattaa lukea. Jos tiivistelmän lukemisen jälkeen aineisto vaikutti vielä hyödylliseltä, perehdyttiin tarkemmin johdantoon, johtopäätöksiin ja tarvittaessa koko aineistoon. Tähän tarkasteluun kuului lisäksi myös aineiston luotettavuuden arviointi. Luotettavuutta arvioitiin tarkastelemalla tutkimuksessa käytettyä kieltä, sen lähteitä ja lähdemerkintöjä sekä aineiston julkaisijaa.

Tutkimusaineiston keräämisessä tärkeässä osassa oli myös haulla löydettyjen julkaisuiden lähdeluetteloiden tarkastelu. Jos julkaisuissa viitattiin hyödylliseltä kuulostavaan aineistoon, viitatuille aineistolle tehtiin aiemmin kuvattu tarkastelu. Tällä tavalla löydettiin aineistoa myös käytännön seulan ulkopuolelta. Näihin aineistoihin oli kuitenkin viitattu useissa käytännön seulaan sopivissa tutkimuksissa, joten ne koettiin tutkimuksen kannalta merkityksellisiksi. Useimmat näin löydettyistä tutkimuksista eivät täyttäneet joko aikarajasta tai ne eivät olleet vertaisarvioituja tutkimuksia.

Myös käytännönläheisemmät aineistot koettiin arvokkaiksi, vaikka ne eivät täysin täyttäneetkään asetettua käytännön seulaa, sillä tämä tutkimus tarkastelee datanhallintaa sen soveltamisen ja käytännön toteuttamisen kannalta. Yksi tällainen käytännön seulan ulkopuolelta löydetty aineisto oli DAMA:n Data Management Body of Knowledge -kirja. Aineisto ei tullut vastaan hauissa, mutta siihen on viitattu useissa käytännön seulaan sopivissa tutkimuksissa. Kirjassa tarkastellaan datanhallintaa käytännönläheisemmästä näkökulmasta, mikä on tutkimuksen kannalta arvokasta, koska kuten edellä mainittu, tutkimus keskittyy datanhallinnan soveltamiseen ja käytännön toteuttamiseen. Näiden

asioiden vuoksi kirja katsottiin työn kannalta merkitykselliseksi aineistoksi. Kirjallisuuskatsauksen lopulliseen aineistoon päätyi 48 lähdettä. Taulukossa 3 on esitelty niistä keskeisimmät sekä niiden lyhyt kuvaus. Taulukon 3 lähteet nostettiin työn keskeisimmäksi aineistoksi, sillä ne todettiin tutkijan tekemän arvion mukaan laadukkaimmiksi ja käytetyimmiksi lähteiksi.

Taulukko 3. *Työn kirjallisuuskatsauksen keskeisin aineisto*

Tekijät	Otsikko	Sisältö
Weber et al. (2009)	One size does not fit all — A contingency approach to data governance	Tutkimuksessa luodaan malli datanhallinnasta ja tunnistaan sen onnistumiseen vaikuttavia tekijöitä. Tutkimus tunnistaa yrityksen kontekstin vaikutuksen datanhallinnan suunnitteluun.
Khatri & Brown (2010)	Designing Data Governance	Artikkelissa luodaan datanhallinnan viitekehys, jonka keskiössä ovat roolit ja päätöksentekoaalueet.
DAMA (2017)	Data Management Body of Knowledge	Kirja on kattava opas datanhallinnan ja datan operatiivisen hallinnan käsitteisiin ja käytäntöihin.
Dallemule & Davenport (2017)	What's Your Data Strategy?	Artikkeli käsittelee datastrategian painottamista datan suojaamisen ja hyödyntämisen välillä.
Mahanti (2018)	Data Governance Implementation: Critical Success Factors	Tutkimuksessa tunnistetaan datanhallinnan toteutuksen kriittisiä menestystekijöitä.
Abraham et al. (2019)	Data governance: A conceptual framework, structured review, and research agenda	Tutkimuksessa luodaan datanhallinnan viitekehys.
Alhassan et al. (2019)	Data governance activities: a comparison between scientific and practice-oriented literature	Tutkimuksessa tunnistetaan kirjallisuudesta datanhallinnan toimintoja ja luodaan niiden avulla malli datanhallinnasta.
Al-Ruithe et al. (2019)	A systematic literature review of data governance and cloud data governance	Tutkimus tunnistaa datanhallinnan sekä pilvidatanhallinnan olennaisia osa-alueita ja haasteita.
Zhang et al. (2022)	Data Matters: A Strategic Action Framework for Data Governance	Tutkimuksessa luodaan viitekehys, jossa yhdistetään datanhallinnan toiminnot sekä niitä tukevat strategiset toimenpiteet.
Walsh. et al. (2022)	Grounding data governance motivations: a review of the literature	Tutkimuksessa tunnistetaan datanhallinnan motivaattoreita.
Lis et al. (2023)	An Investigation of Antecedents for Data Governance Adoption in the Rail Industry-Findings From a Case Study at Thales	Tutkimuksessa tunnistetaan datanhallinnan omaksumiseen vaikuttavia tekijöitä.

Työn keskeisin aineisto koostuu pääosin erilaisista datanhallinnan viitekehyksistä ja maleista. Työssä hyödynnettiin aineistoa, joissa datanhallintaa tarkastellaan ihmisten, toimintojen ja prosessien näkökulmasta. Tutkimuksesta jätettiin pois lähteitä, jotka käsitelivät datanhallintaa pääasiassa erilaisten järjestelmien tai teknologioiden, kuten lohko-
ketjujen kautta. Myös tutkimuksen konteksti saattoi vaikuttaa aineiston poisjäämiseen.

Esimerkiksi jotkin julkiseen sektoriin tai suurempiin ekosysteemeihin liittyvät datanhallinnan viitekehykset jätettiin tutkimuksesta pois, koska niiden koettiin tarkastelevan aihetta tutkimuksen kannalta epärelevantista näkökulmasta.

2.3 Tapaustutkimus

Tämän tutkimuksen empiirinen osuus toteutettiin holistisena yksittäisen tapauksen tapaustutkimuksena. Bexter ja Jack (2008) esittävät, että holistinen yksittäisen tapauksen tapaustutkimus sopii tilanteisiin, joissa tutkitaan yhtä ympäristöä sen uniikkiuden tai äärimmäisyyden vuoksi. Tässä tutkimuksessa uniikkina ympäristönä toimii tutkimuskohteenä oleva organisaatio ja tutkimuksessa tutkitaan yhtä tapausta, kohdeorganisaation datanhallintaa. Vuoren (2021b) mukaan tapaustutkimuksessa on tärkeää liittää tapaus kiinteästi kontekstiinsa. Alla on kuvattu tarkemmin kohdeorganisaation datanhallinnan kontekstia kohdeorganisaation esittelyn kautta.

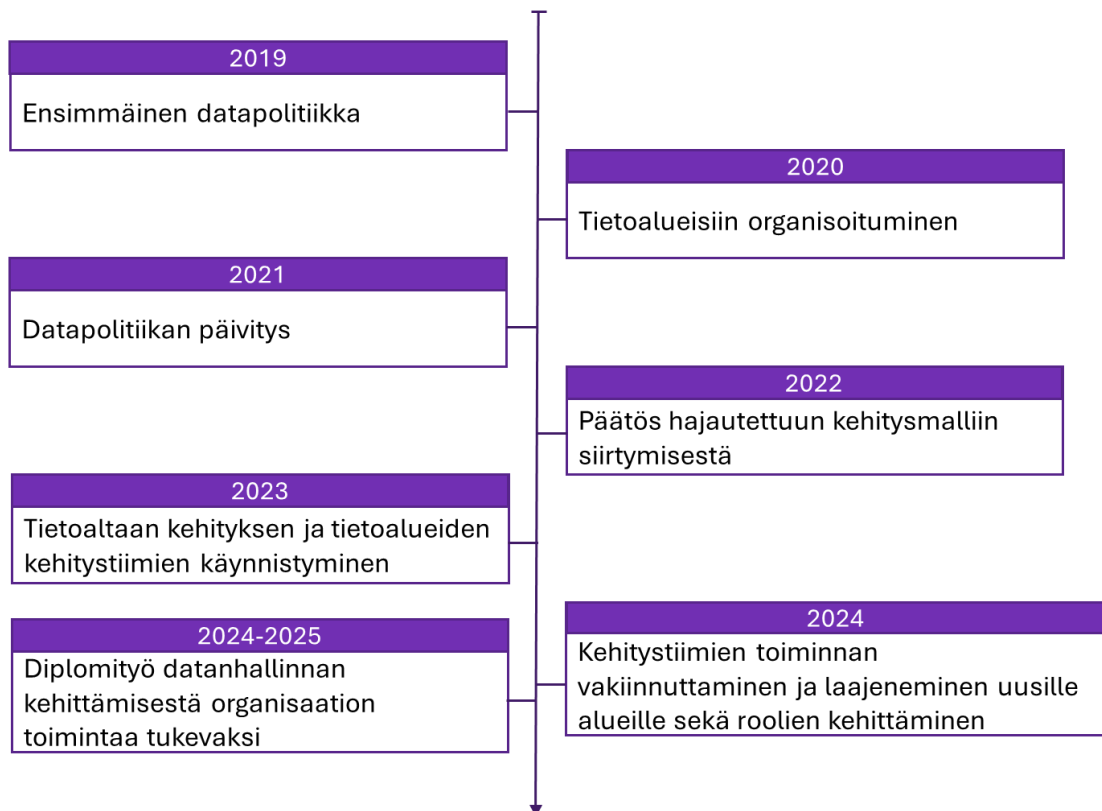
2.3.1 Kohdeorganisaation esittely

Tutkimuksen kohdeorganisaationa toimii suomalainen kantaverkkoyhtiö Fingrid Oyj. Fingrid vastaa maanlaajuisesta kantaverkosta, eli sähkönsiirron runkoverkosta, johon liittyvät suuret sähköntuottajat, paljon sähköä kuluttavat tehtaot sekä sähkön jakeluverkot. Fingridin tehtävä on turvata varma sähkö Suomessa kaikissa tilanteissa sekä edistää puhdasta, markkinaehtoista sähköjärjestelmää. Yhtiön omistaa Suomen valtio ja suomalaiset eläkeyhtiöt. Fingridillä on luonnollinen monopoli, joka perustuu siihen, että ei ole kansantaloudellisesti järkevää rakentaa kallista päällekkäistä infrastruktuuria useiden yhtiöiden toimesta. Vuonna 2024 Fingridin henkilöstön koko oli 597 ja sen liikevaihto oli noin 1 269 miljoonaa euroa. (Fingrid)

Kohdeorganisaation toimintaympäristössä vaikuttavat megatrendit, ilmastonmuutos ja digitalisaatio, näkyvät vahvasti yhtiön toiminnassa. Ilmastonmuutoksen torjunta ja siirtyminen vihreämpään sähköön aiheuttaa muutoksia suomalaisessa sähköjärjestelmässä ja asettaa uusia vaatimuksia yhtiön toiminnalle. Digitalisaation myötä myös datan määrän kasvu ja sen laadun hallinta haastaa yritystä. (Fingrid-lehti 2019; Fingrid-lehti 2024)

Fingrid valittiin tutkimuksen kohteeksi, koska yhtiöllä on käytössä datanhallintamalli, jota se pyrkii aktiivisesti kehittämään. Tutkimuksessa tarkastellaan datanhallintaa kokonaisuutena, sen suunnittelun ja toteuttamisen kautta, joten käytössä oleva datanhallintamalli oli olennainen kriteeri. Lisäksi Fingridillä on tunnistettu valmiiden datanhallintamallien soveltamisen vaikeus sekä tarve selvittää datanhallinnan tilaansa, joten tutkimuksella pystytään tuottamaan arvoa myös kohdeorganisaatiolle. Fingridin valintaa puolsi myös se, että tutkijalla oli pääsy yhtiötä koskevaan aineistoon.

Viime vuosina Fingridillä on tehty useita toimenpiteitä datanhallinnan kehittämiseksi. Näistä keskeisimmät on esitelty kuvassa 3. Yksi datanhallinnan keskeinen kehityssuunta on ollut sen liiketoimintalähtöisyyden lisääminen. Organisaatio on jaettu tietoalueisiin, jotka edustavat liiketoiminnan osa-alueita tai prosesseja. Samalla keskitetystä raportti- ja datakehityksestä on siirrytty hajautettuun malliin. Hajautetussa mallissa vastuu kehityksestä on jaettu liiketoiminnan osa-alueille muodostamalla tietoalueille omat tiimit data- ja analytiikkakehitykseen. Tämän tavoitteena on ollut tuoda data- ja analytiikkaresurssit lähemmäs liiketoimintaa. Organisaatiossa on myös otettu käyttöön uusia datanhallinnan rooleja, kuten tietoalueen omistaja, sekä muodostettu datanhallintaryhmä, joka vastaa datahallinnan kehityksestä koko organisaation tasolla. Näiden lisäksi yhtiössä on aloitettu uuden tietoaaltaan käyttöönotto, mikä on vaikuttanut viime aikoina paljon datakehitykseen, raportointiin sekä toimintamallien kehitykseen. Kuten kuvasta 3 nähdään, kohdeorganisaation datanhallinta on vahvasti yhteydessä yhtiön data- ja analytiikkakehitykseen, minkä vuoksi myös tässä työssä datanhallintaa tarkastellaan paljon kyseisen näkökulman kautta. Tämä on perusteltua, sillä datanhallinnan linjaukset näkyvät erityisesti data- ja analytiikkakehityksessä ja siihen osallistuvien ihmisten työssä.



Kuva 3. Kohdeorganisaation datanhallinnan kehityksen merkittävimmät tapahtumat sekä tutkimuksen toteutus aikajanalla

Kuvasta 3 nähdään, että vuonna 2024 kohdeorganisaatiossa laajennettiin kehitystiimien toimintaa ja kehitettiin silloisia datanhallinnan rooleja. Kuten johdannossa todettiin, organisaatiossa ei haluttu ottaa käyttöön valmiita roolituksia, jotka koettaisiin liian byrokraattisiksi tai hankaliksi, joten organisaatiossa syntyi tarve löytää keinoja uudistaa datanhallintamallia kohdeorganisaation toimintaan sopivaksi. Tämä loi tarpeen tämän tutkimuksen toteuttamiselle. Lisäksi roolien tarkastelu nosti esiin työn taustalla olevan tarpeen datanhallinnan nykytilan selvittämisestä eli tarpeen tarkastella laajemminkin kohdeorganisaation datanhallinnan tilaa ja sopivuutta organisaation toimintaan. Kuvasta nähdään, että roolien kehitys alkoi saman vuonna kuin tutkimuksen toteutus. Roolien kehitys oli osittain vielä kesken tutkimuksen toteutuksen aikaan, joten se on mahdollisesti voinut vaikuttaa haastateltavien vastauksiin.

2.3.2 Empiirisen aineiston keruu

Tutkimuksen empiirisen osan varsinaisena aineistonkeruumenetelmänä toimi puolistrukturoidut haastattelut. Ennen haastattelua haastateltaville lähetettiin kuitenkin taustoittava kysely, johon he vastasivat ennen haastattelutilannetta. Kyselyn tarkoituksena oli helposti vertailtavien ja koostettavien tulosten saamisen lisäksi valmistella vastaajia haastatteluun sekä johdatella aiheeseen. Kyselyn keskeisimpiä osia käytiin haastattelussa läpi ja lisäkysymyksillä pyrittiin keräämään syvällisempää tietoa, kuten perusteluja ja esimerkkejä.

Taustoittava kysely muodostettiin työn teoreettisen viitekehyksen sekä kohdeorganisaation sisäisen datanhallintaa koskevan strategiadokumentin perusteella. Strategiadokumentissa on määritelty kohdeorganisaation datanhallintamallin rakenne sekä datanhallinnan visio, joka koostuu kolmesta osasta: data löytyy luotettavasti, datan hyödyntäminen on helppoa ja dataratkaisujen tekeminen on nopeaa. Dokumentissa on lisäksi kuvattu datanhallinnan visioon liittyen kolme pilaria, jotka kuvaavat erilaisia tapoja hyödyntää dataa ja analytiikkaa tehokkaasti liiketoiminnan tavoitteiden saavuttamisessa. Pilarit ovat: toiminnan seuraaminen ja raportointi, toiminnan optimointi edistyneen analytiikan avulla sekä datan ja analytiikan hyödyntäminen transformatiivisten toimintamallimuutosten edistämässä. Näiden pohjalta taustoittavan kyselyn osa-alueiksi muodostui vastaajan taustatiedot, datanhallinnan merkitys ja tavoitteet, datanhallinnan selkeys ja soveltuvuus, datanhallinnan toteutus sekä datanhallinnan vaikutukset. Taustoittavan kyselyn kysymykset on esitelty liitteessä A.

Tutkimuksen varsinaiseksi aineistonkeruumenetelmäksi valittiin haastattelut niiden joustavuuden takia. Toisin kuin kyselyssä, haastattelijalla on mahdollisuus oikaista väärinkä-

sityksiä, tarkentaa kysymyksiä ja käydä keskustelua haastateltavan kanssa. Haastattelun etuna voidaan pitää myös mahdollisuutta valita haastateltaviksi henkilöitä, joilla on kokemusta tai tietoa tutkittavasta aiheesta. (Tuomi & Sarajärvi 2018) Näiden etujen lisäksi työn tutkimusfilosofia, interpretivismi, ja tutkimusmetodologia, laadullinen tutkimus, tukevat menetelmän valintaa. Työn tutkimusfilosofiassa korostuu eri ihmisryhmien näkemykset ja kokemukset ja laadullisessa tutkimuksessa ei-numeerisen datan merkitys. Tutkimusmenetelmänä haastattelut tukevat näitä molempia.

Tuomen ja Sarajärven (2018) mukaan haasteena haastatteluissa voidaan pitää niiden aikaa vievyyttä. Toisaalta nopeammin toteutettavissa kyselyissä vastaamattomuus saattaa muodostua ongelmaksi. Laadullisen tutkimuksen haastatteluissa näin ei ole, koska laadullisessa tutkimuksessa ei ole tarkoitus tehdä laajoja yleistyksiä. Jos laadullisen tutkimuksen aineisto jää niukaksi, voidaan kuitenkin kyseenalaistaa, kertooko tutkimus mitään tutkittavasta aiheesta. Tutkijoiden mukaan tämä voi tapahtua, jos haastateltavilla ei ole tarpeeksi kokemusta tutkittavasta asiasta.

Haastattelut toteutettiin puolistrukturoituina haastatteluina, mikä tarkoittaa sitä, että tutkijalla on valmiit kysymykset, joihin haastateltava saa vastata vapaasti (Hyvärinen et al. 2021). Puolistrukturoitujen haastatteluiden yksi merkittävimmistä eduista on niiden mahdollistama vuorovaikutus haastattelijan ja osallistujan välillä. Ne mahdollistavat tutkijalle jatkokysymysten tekemisen haastateltavan vastausten perusteella sekä antavat tilaa haastateltavien yksilölliselle sanalliselle ilmaisulle. (Kallio et al. 2016) Puolistrukturoitujen haastatteluiden käyttäminen vaatii tutkimusaiheeseen liittyvään kirjallisuuteen ja tutkimukseen perehtymistä, koska haastattelukysymykset pohjautuvat aiempaan tutkimukseen (Kallio et al. 2016). Tässäkin tutkimuksessa haastattelurunko muodostettiin pääasiassa tehdyn kirjallisuuskatsauksen perusteella.

Haastatteluiden tavoitteena oli saada perusteluita taustoittavan kyselyn vastauksille sekä kerätä syvällisempää tietoa tutkimuksen aihepiireistä. Haastattelurunko sisälsi kysymyksiä, joissa viitattiin suoraan haastateltavan taustoittavan kyselyn vastauksiin. Lisäksi haastattelurungossa oli kysymyksiä datanhallinnan tilaan, rooleihin ja haasteisiin liittyen, sillä haastattelun sisältöä ei haluttu rajoittaa vain kyselyn vastausten perusteluun ja avaamiseen. Muuten kyselyn perustana ollut teoreettinen viitekehys olisi todennäköisesti ohjannut ja rajoittanut liikaa viitekehysten ulkopuolisten näkökulmien esiin nousemista. Haastattelurunko on kuvattu liitteessä B.

Kuten aiemmin todettu, laadulliselle tutkimukselle on tyypillistä harkinnanvarainen otanta (Eskola & Suoranta 1998). Myös tässä tutkimuksessa haastateltavat valittiin harkinnanvaraisesti eikä sattumanvaraisesti. Koska laadullisen tutkimuksen otanta on pieni, ei ole

järkevää valita tutkittavaa joukkoa sattumanvaraisesti vaan tutkittaviksi kannattaa valita henkilöitä, jotka kuuluvat tutkimuksen kohderyhmään (Eskola & Suoranta 1998). Näin varmistetaan, että tutkimuksen kohteilta saadaan tutkimuksen kannalta relevanttia tietoa. Haastateltaviksi pyrittiin valitsemaan henkilöitä, joilla oli kokemusta datanhallinnasta tai data- ja raporttikehityksestä joko data-asiantuntijan tai liiketoiminnan asiantuntijan näkökulmasta. Harkinnanvaraisella otannalla varmistettiin myös, että tietoa saatiin kattavasti eri organisaatiotasoilta. Seuraavassa luvussa esitellyn teoreettisen viitekehyksen perusteella datanhallinnan roolit voidaan sijoittaa neljälle eri organisaatiotasolle: strateginen ohjaus, strateginen toteutus, taktinen sekä operatiivinen (kuva 6). Jotta nämä tasot tulivat edustetuiksi, kahdesta Fingridin yksiköstä valittiin haastateltaviksi ylintä johtoa, päälliköitä, kehitystiimin vetäjiä sekä asiantuntijoita. Osa haastateltavista oli hiljattain vaihtanut toimintoa organisaation sisällä, joten tulokset eivät kuitenkaan rajaudu tiukasti vain kahteen yksikköön. Osalla haastateltavista oli myös kokemusta useammalla organisaatiotasolla toimimisesta, mikä voi vaikuttaa heidän vastauksiinsa. Haastateltavat on esitetty taulukossa 4.

Taulukko 4. *Haastateltavat*

Haastateltava	Toimenkuva	Haastattelun kesto
H1	Kehitystiimin vetäjä	51 min
H2	Kehitystiimin vetäjä	67 min
H3	Asiantuntija	46 min
H4	Asiantuntija	52 min
H5	Päällikkö	49 min
H6	Ylin johto	48 min
H7	Asiantuntija	49min
H8	Kehitystiimin vetäjä	36 min
H9	Päällikkö	35 min
H10	Päällikkö	61 min

Taustoittava kysely lähetettiin haastateltaville noin viikko ennen haastattelua, jotta vastaajilla oli tarpeeksi aika vastata kyselyyn ja tutkijalla oli mahdollisuus katsoa kyselyn vastaukset läpi ennen haastattelutilannetta. Tämä oli oleellista, koska haastattelutilanteessa haluttiin syventyä taustoittavassa kyselyssä esiinnousseisiin asioihin. Itse haastattelukysymyksiä ei lähetetty etukäteen, koska kyselyn nähtiin toimivan tarpeellisena valmistautumisena haastatteluun, eikä vastaajia ei haluttu kuormittaa liialla ennakkomateriaalilla. Ennen haastattelua vastaajilta tiedusteltiin halukkuutta osallistua haastatteluun, lähetettiin kalenterikutsu ja myöhemmin vielä taustoittava kysely, joten haastattelukysymysten lähettäminen ja niihin perehtyminen etukäteen olisi voitu kokea kuormittavaksi. Haastattelut pidettiin aikavälillä 18.11.2024-12.12.2024. Haastattelut toteutettiin paikan päällä tai etänä Microsoft Teamsin välityksellä.

2.3.3 Empiirisen aineiston analyysi

Taustoittavan kyselyn tuloksista tehtiin visualisointeja ja sen pääkohdat avattiin sanallisesti, mutta tuloksia ei analysoitu tilastollisin menetelmin pienen otoksen vuoksi. Taustoittavan kyselyn tulokset on esitelty luvussa 4.1. Haastattelut sen sijaan analysoitiin teoriaohjaavan sisällönanalyysin keinoin. Teoriaohjaavassa analyysissä teoria toimii analyysin apuna, mutta analyysi ei perustu siihen suoraan. Teoriaohjaavassa analyysissä analyysiyksiköt valitaan aineistosta, mutta aikaisempi tieto auttaa ja ohjaa analyysin tekemistä. Analyysissä tunnistetaan aikaisemman tiedon vaikutus, mutta sen tarkoitus ei ole puhtaasti testata teoriaa, toisinkuin teorialähtöisessä analyysissä usein on. (Tuomi & Sarajärvi 2018) Tämän työn haastatteluaineiston analyysissä teemat muodostettiin aineistolähtöisesti, mutta kirjallisuuskatsauksen avulla muodostettu teoreettinen viitekehys auttoi niiden tunnistamisessa. Teoria ei kuitenkaan rajoittanut analyysia ja aineistosta pyrittiin tunnistamaan myös uusia, viitekehyksessä tunnistamattomia teemoja. Teoriaohjaava analyysi valittiin, sillä teoreettista viitekehystä halutaan testata, mutta myös jatkojalostaa peilaamalla lopuksi empiirisen tutkimuksen löydöksiä työn teoreettiseen viitekehukseen.

Laadullisen tutkimuksen sisällönanalyysi seuraa usein alla kuvattua prosessia.

1. Vahva päätös siitä, mikä aineistossa on kiinnostavaa
2. Aineiston läpikäyminen ja kiinnostuksen kohteeseen liittyvien asioiden erottaminen ja merkitseminen sekä näihin kuulumattomien asioiden poisjättäminen
3. Merkittyjen asioiden yhteen kerääminen erilleen muusta aineistosta
4. Aineiston luokittelu, teemoittelu tai tyypittely
5. Yhteenvedon kirjoittaminen. (Tuomi & Sarajärvi 2018)

Haastatteluaineiston analyysin ensimmäisen vaiheen määrittelee tutkimuksen tarkoitus tai tutkimusongelma. Niiden tulee olla linjassa aineiston analyysin mielenkiinnonkohteen kanssa. (Tuomi & Sarajärvi 2018) Tässä tutkimuksessa analyysissä keskityttiin kohdeorganisaation datanhallintamalliin, datanhallinnan toteuttamiseen sekä näiden haasteisiin ja kehityskohteisiin.

Analyysiprosessin toinen ja kolmas kohta tarkoittavat haastatteluiden litterointia eli tekstiksi muuttamista sekä koodaamista. Litteroidut haastattelut muodostavat haastattelututkimuksen tutkimusaineiston. Litterointi on tärkeä vaihe aineiston analyysin kannalta. Sen avulla tutustutaan aineistoon ja otetaan etäisyyttä haastattelutilanteeseen. Litteroinnin

tarkkuus riippuu tutkimuskysymyksistä ja analyysitavasta. Litteroinnissa voidaan keskittyä pelkästään puheen sisältöön tai siinä voidaan huomioida myös tapa, jolla asiat sanotaan. (Hyvärinen et al. 2017) Tässä tutkimuksessa haastattelut tallennettiin ja litteroitiin automaattisesti Teamsin transcript-toiminnon avulla. Toimintoa voitiin soveltaa haastatteluiden litteroinnissa, koska analyysissa haluttiin keskittyä siihen, mitä haastateltavat sanovat. Tällaisessa tutkimuksessa haastatteluista litteroidaan vain puheen sisältö, mihin Teamsin automaattinen transcript-toiminto soveltui hyvin. Haastatteluiden jälkeen automaattinen litterointi tarkistettiin ja siihen tehtiin korjauksia, sillä automaatio tulkitse sanoja välillä väärin. Litteroinnista poistettiin myös tutkimuksen aiheisiin liittymättömiä osia. Paikan päällä pidettyjen haastatteluiden litteroinnista täytyi myös erotella, kuka puhui ja mitä, sillä Teams tulkitse samassa tilassa olevat henkilöt yhdeksi puhujaksi.

Aineiston koodaaminen on sisällönanalyysin yksi työväline, jonka avulla aineiston käsittely aloitetaan. Se on aineiston alkuvaiheen järjestämistä ja luokittelua, minkä jälkeen toteutetaan varsinainen aineiston analyysi. Koodaaminen tarkoittaa käytännössä aineiston eri osien yhdistelyä ja erottelua jonkin ominaisuuden mukaan. (Juhila 2021a) Tässä työssä koodaus toteutettiin käsin Microsoft Wordin ja Excelin avulla. Litterointeja käsiteltiin Wordissa, jossa tiettyyn aihepiiriin liittyvät asiat korostettiin samalla värillä ja väri ja luokkaa yhdistävä tekijä kirjattiin Exceliin ylös. Exceliin kirjattiin myös luokkaa kuvaavia litteroinnin osia. Tämän jälkeen koodatut haastatteluaineistot ryhmiteltiin haastateltavan roolin mukaan. Ryhmittelyssä käytettiin aikaisemmin esiteltyä jakoa ylimpään johtoon, päälliköihin, kehitystiimin vetäjiin ja asiantuntijoihin. Jako tehtiin haastateltavan nykyisen roolin perusteella.

Työn sisällönanalyysin neljännessä vaiheessa hyödynnettiin teemoittelua. Teemoittelu on laadullisen aineiston pilkkomista ja ryhmittelyä erilaisten aihepiirien mukaan. Sen tarkoituksena on mahdollistaa tiettyjen teemojen esiintymisen vertailu aineistossa. (Tuomi & Sarajärvi 2018) Teemoittelussa tulee huomioida, että aineiston analyysissa esiin nousevat teemat eivät ole sama asia kuin haastattelun teemat. Jos tutkija jäsentää keräämänsä aineiston haastattelun pääteemojen mukaan, tutkijan ennakkokäsitys päättyy ohjaamaan analyysia. Tämä vääristää analyysin tuloksia, sillä tutkijan tulisi mahdollistaa erilaisten teemojen ja jäsenysperiaatteiden nouseminen aineistosta. (Juhila 2021b) Tässä tutkimuksessa teemoittelu toteutettiin käyttäen pohjana koodauksen avulla luotuja luokkia, joita yhdistettiin tutkimusongelman kannalta relevanteiksi teemoiksi ja pääteemoiksi eli datasta tunnistettiin yhteneväisiä merkityksiä. Teemojen tunnistamisen ja nimeämisen apuna hyödynnettiin työn teoreettista viitekehystä, mutta se ei kuitenkaan määrännyt, mitä teemoja aineistosta tunnistettiin. Aineistossa mainittiin esimerkiksi joustavuus jossain muodossa useammassa yhteydessä. Joustavuudesta olisi voitu luoda

teema, mutta koska työn teoreettinen viitekehys vaikutti analyysin taustalla, maininnat jaettiin teemoiksi sen mukaan, missä yhteydessä joustavuus mainittiin. Esimerkiksi eräässä haastatteluissa todettiin, että toimintaympäristön nopean muuttumisen vuoksi tarvitaan joustavia toimintamalleja. Koska viitekehyksessä tunnistettiin asioita, jotka vaikuttavat datanhallintamallin muodostamiseen, maininta laitettiin toimintaympäristön nopea muuttuminen -teeman alle eikä sitä luotu teemaa ”joustavuus”. Alustavan teemoittelun jälkeen aineistoa käytiin läpi ja teemoja arvioitiin, yhdisteltiin ja hiottiin. Niiden esiintymistä tarkasteltiin myös eri organisaatiotasolla. Teemoja peilattiin tutkimuskysymyksiin ja tutkimuksen tavoitteisiin, jotta varmistettiin, että ne ovat tutkimuksen kannalta relevantteja.

Sisällönanalyysissa tutkijan tulisi pystyä kuvaamaan, mikä sisällöissä on kiinnostavaa, miten aineiston eri osat eroavat ja mikä niitä yhdistää sekä miten sisältöjen yksityiskohdista voi muodostaa kokonaiskuvan. Pelkkä aineiston sisällön kuvaus ei siis ole sisällönanalyysia. (Vuori 2021a) Tässä työssä huomioita kirjoitettiin jatkuvasti ylös analyysin edetessä, mutta varsinainen yhteenveto ja johtopäätökset, eli analyysiprosessin kohta 5, kirjoitettiin vasta lopullisten teemojen muodostamisen jälkeen.

3. TEOREETTINEN VIITEKEHYS

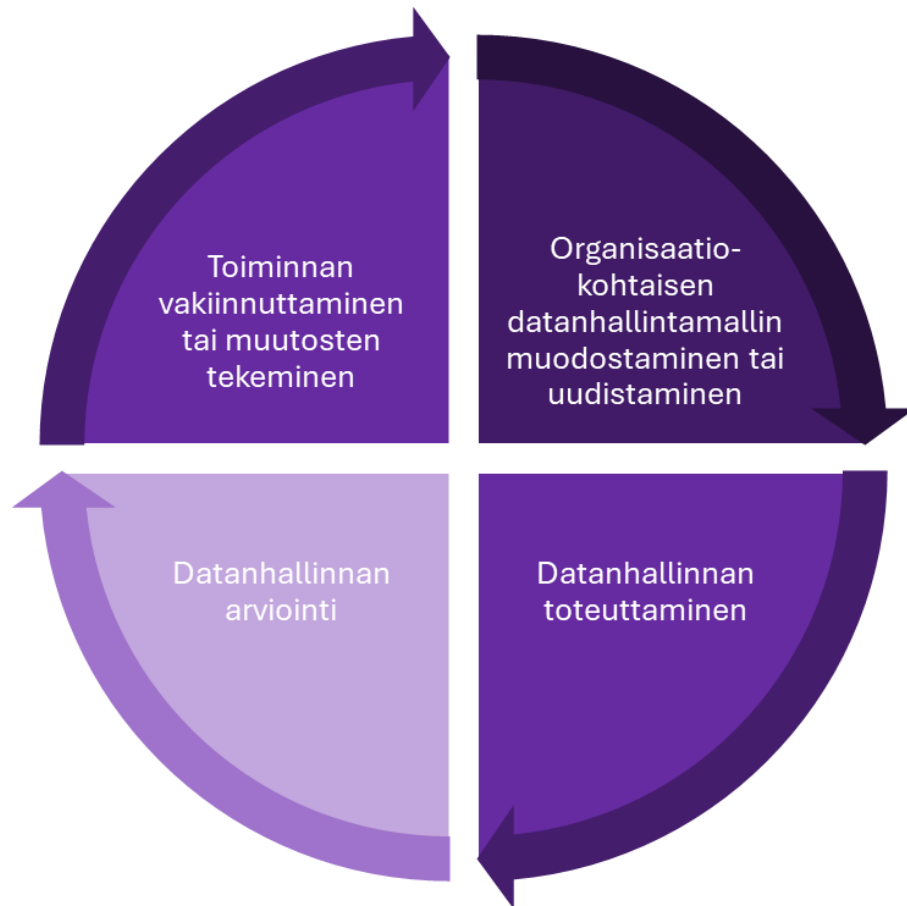
Tässä luvussa rakennetaan tutkimukselle teoreettinen viitekehys. Työn teoreettinen viitekehys pohjautuu näkemykseen, jonka mukaan datanhallintaa ei tulisi pitää kertaluontoisena projektina vaan jatkuvana prosessina (DAMA 2017; Walsh et al. 2022), jota pitää pystyä sopeuttamaan ulkoisiin ja sisäisiin muutoksiin (DAMA 2017; Abraham et al. 2019; Mahanti 2019). Tämän vuoksi PDCA-malli, eli jatkuvan parantamisen Plan-Do-Check-Act-malli, valittiin teoreettisen viitekehyyksen perustaksi.

PDCA-malli on Demingin laajaan käyttöön vakiinnuttamaa iteratiivinen neljän vaiheen johtamismalli, jota käytetään prosessien ja tuotteiden jatkuvaan parantamiseen. Mallin vaiheet ovat suunnittele (plan), tee (do), tarkista (check) ja toimi (act). (Deming 1986, lähteestä Peng et al. 2018) Vaiheet kuvataan usein ympyrässä, sillä jatkuvaan parantamiseen kuuluu vaiheiden toistaminen (Isniah et al. 2020). Mallin ensimmäisessä vaiheessa asetetaan tavoitteet ja suunnitellaan toimenpiteet niiden saavuttamiseksi. Toisessa vaiheessa toteutetaan suunnitelma. Kolmannessa vaiheessa arvioidaan vaikutuksia ja viimeisessä vaiheessa toimenpiteet vakiinnutetaan tai niihin tehdään muutoksia. (Peng et al. 2018; Isniah et al. 2020)

PDCA-malli auttaa organisaatioita suhtautumaan datanhallintaan jatkuvana prosessina. Malli mahdollistaa myös datanhallinnan toimivuuden jatkuvan arvioinnin ja reagoimisen ulkoisiin ja sisäisiin muutoksiin. Jos datanhallintaa kehitetään vain kertaluontoisesti organisaation toimintaan sopivaksi, sen toimivuus todennäköisesti heikkenee ajan kuluessa, sillä harvan organisaation tavoitteet ja toimintaympäristöt ovat muuttumattomia. PDCA-mallin valintaa puolsi myös se, että sitä voidaan käyttää jo käyttöönotettujen prosessien arviointiin ja parantamiseen sekä täysin uusien prosessien luomiseen (Peng et al. 2018; Isniah et al. 2020). Työssä halutaan luoda datanhallinnan kehitysmalli, jota voidaan hyödyntää tapauksissa, joissa organisaatio haluaa jatkokehittää datanhallintamalliaan sekä tapauksissa, joissa organisaatio vasta muodostaa ensimmäistä datanhallintamalliaan. PDCA-mallin valinta viitekehyyksen pohjaksi mahdollistaa näiden molempien tapausten huomioimisen. Lisäksi mallia hyödynnetään kohdeorganisaatiossa jo muissa yhteyksissä, joten se on helppo ymmärtää ja omaksua.

PDCA-mallin hyödyntäminen datanhallinnan tutkimuksessa ei ole täysin uusi idea, sillä sitä on aiemmin hyödynnetty myös muissa datanhallintaa käsittelevissä tutkimuksissa (Peng et al. 2018; Maulina & Ruldeviyani 2019; Romero et al. 2019). Tässä tutkimuksessa PDCA-mallia ei hyödynnetä sellaisenaan, vaan sitä muokataan datanhallintaan

sopivaksi ja se yhdistetään muihin teorioihin. PDCA-mallia soveltaen teoreettisen viitekehysten rakenteeksi määritellään organisaatiokohtaisen datanhallintamallin muodostaminen tai uudistaminen (plan), datanhallinnan toteuttaminen (do), datanhallinnan arviointi (check) sekä toiminnan vakiinnuttaminen tai muutosten tekeminen (act). Nämä vaiheet luovat pohjan työn teoreettiselle viitekehykselle. Sovellatut vaiheet on esitetty kuvassa 4.



Kuva 4. PDCA-viitekehys sovellettuna organisaatiokohtaisen datanhallinnan kehittämiseen (mukaillen Deming 1986, lähteestä Peng et al. 2018)

Viitekehysten rakentaminen alkaa datanhallinnan määrittelyllä ja datanhallintamallien osa-alueiden esittelyllä. Tämän jälkeen kuvataan organisaatiokohtaisen datanhallintamallin suunnittelussa tai uudistamisessa huomioitavat tekijät, joiden keskiössä ovat organisaation tavoitteet sekä kontekstitekijät. Organisaatiokohtaisen datanhallintamallin suunnittelun jälkeen tarkastellaan datanhallinnan toteuttamista ja siihen vaikuttavia tekijöitä. Tämän jälkeen tarkastellaan datanhallinnan arviointia. Toiminnan vakiinnuttaminen tai muutosten tekeminen jää teoriaosasta pois, koska se liittyy enemmän käytännön toteuttamiseen ja riippuu syklin muista vaiheista. Viimeisenä teoriaosasta tehdään vielä yhteenveto ja synteesi.

3.1 Datanhallinta yleisesti

Datanhallinnalle ei ole olemassa täysin vakiintunutta määritelmää (Otto 2011a; Abraham et al. 2019; Al-Ruithe et al. 2019; Walsh et al. 2022). Tästä huolimatta eri määritelmistä voidaan tunnistaa yhteisiä piirteitä. Useita datanhallinnan määritelmiä yhdistää näkemys datasta strategisena resurssina, jonka arvoa pyritään kasvattamaan (Khatri & Brown 2010; Otto 2011a; Zhang et al. 2022). Tähän pyritään määrittelemällä hallintamekanismeja, kuten rakenteita, kontrollimekanismeja ja erityisesti rooleja ja vastuita, joilla varmistetaan datan asianmukaisen käyttö ja tuetaan päätöksentekoa (Khatri & Brown 2010; Otto 2011a; DAMA 2017; Abraham et al. 2019; Karkošková 2023). Yksinkertaistettuna datanhallinta voidaan tulkita datan käytön ja hallinnan mahdollistavaksi organisoitumismalliksi. Kuten aiemmin mainittu, datanhallinta ei kuitenkaan ole kertaluontoisesti toteutettava projekti, vaan sitä voidaan pitää jatkuvana prosessina, joka vaikuttaa organisaation datan lisäksi ihmisiin, prosesseihin ja teknologiaan (Walsh et al. 2022).

Yhden yleisesti hyödynnetyn datanhallinnan määritelmän tarjoaa Data Management Association, DAMA. DAMA (2017) määrittelee datanhallinnan auktoriteetin ja kontrollin (suunnittelu, seuranta ja täytäntöönpano) käyttönä dataomaisuuksien hallinnassa. DAMA:n mukaan tähän kuuluu muun muassa datan hallintaan liittyvien käytänteiden, vastuiden ja mittareiden määrittäminen, kommunikointi ja käyttöönotto. DAMA:n mukaan datanhallinta varmistaa, että dataa hallitaan asianmukaisesti, linjausten ja parhaiden käytänteiden mukaan.

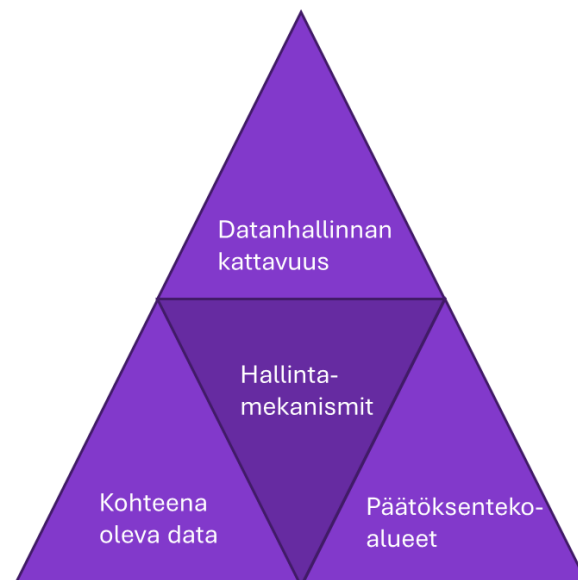
Datanhallinnan käsite liittyy läheisesti datan operatiiviseen hallintaan (engl. data management). Datan operatiivinen hallinta sisältää dataan liittyvien toimintojen käytännön toteuttamisen (Karkošková 2023). Näitä toimintoja ovat muun muassa datan tallennus, integraatiot, datan varastointi, liiketoimintatiedonhallinta ja datan laadunhallinta (DAMA 2017). Kun datanhallinta viittaa siihen, mitä päätöksiä pitää tehdä ja kuka tekee nuo päätökset, datan operatiivinen hallinta viittaa näiden päätösten tekemiseen osana päivittäistä datapolitiikkojen toteuttamista (Abraham et al. 2019). Datanhallinnan voidaan nähdä toteuttavan valvontaa ja varmistavan, että dataa hallitaan, kun taas datan operatiivinen hallinta toteuttaa tätä hallintaa määriteltyjen tavoitteiden saavuttamiseksi (DAMA 2017). Voidaankin siis todeta, että datanhallinta ohjaa datan operatiivisen hallinnan toimintoja. Se voidaan käsittää malliksi, jonka avulla datan operatiivinen hallinta organisoidaan.

Edellä esitellystä datanhallinnan kuvauksesta nähdään, että datanhallinnan keskiössä on luoda viralliset raamit ja rakenne organisaation datankäsittelylle erilaisten hallintame-

kanismien, kuten roolitusten ja käytäntöjen, kautta. Datanhallinnan tavoitteena on varmistaa datan kohtelemisen arvokkaana resurssina ja mahdollistaa datan asianmukainen hyödyntäminen ja siitä arvon luominen.

3.2 Organisaatiokohtaisen datanhallintamallin muodostaminen

Datanhallinta tulee suunnitella organisaatiokohtaisesti, koska kuten jo johdannossa todettiin, ei ole olemassa yhtä oikeaa ja kaikille organisaatioille sopivaa tapaa järjestää datanhallinta (Weber et al. 2009; Al-Ruithe & Benkhelifa 2017). Tässä työssä organisaatiokohtaisen datanhallinnan kehittämistä tarkastellaan ensiksi datanhallintamallien kautta. Datanhallintamallit määrittelevät yleensä minkälaista dataa halutaan hallita (DAMA 2017; Abraham et al. 2019), mitä hallintaan liittyviä päätöksiä täytyy tehdä (Weber et al. 2009; Khatri & Brown 2010), millä alueilla (Khatri & Brown 2010; Abraham et al. 2023) ja kuka tekee päätökset, eli mitä päätöksentekorooleja tarvitaan (Weber et al. 2009; Otto 2011a). Tässä työssä datanhallintamallin rakenteen pohjana käytetään Abrahamin ja muiden (2019) kehittämää viitekehystä (kuva 5). Se jakaa datanhallinnan hallintamekanismeihin ja hallintamekanismien sovelluskohteisiin, joita ovat datanhallinnan kattavuus, kohteena oleva data sekä päätöksentekoalueet. Abrahamin ja muiden viitekehysten hyödyntämistä puoltaa sen kattavuus ja tunnettuus. Viitekehystä on hyödynnetty useissa tutkimuksissa (Ahmadi 2022; Karkošková 2023; Lis 2023; Fattah 2024). Pelkästään tietokanta Scopuksesta löytyy yli 300 tutkimusta, joissa siihen on viitattu.



Kuva 5. Datanhallinnan osa-alueet (mukaillen Abraham et al. 2019)

Abrahamin ja muiden (2019) mukaan datanhallinnan hallintamekanismit koostuvat virallisista rakenteista, jotka yhdistävät liiketoiminnan, tietotekniikan ja datan operatiivisen

hallinnan toiminnot. Lisäksi niihin kuuluvat viralliset päätöksenteko- ja seurantaprosessit ja käytännöt. Tutkijoiden mukaan nämä mekanismit kattavat menettelylliset sekä rakenteelliset hallintamekanismit. Menettelyllisiin hallintamekanismeihin sisältyy muun muassa datastrategia, datapolitiikat, -standardit, -prosessit sekä menettelytavat. Sen sijaan rakenteellisiin hallintamekanismeihin sisältyy roolit ja vastuut sekä päätöksenteko-oikeuksien allokointi. Abraham ja muut sisällyttävät hallintamekanismeihin myös sidosryhmien osallistumista ja yhteistyötä edistävät käytännöt, mutta tässä tutkimuksessa niitä tarkastellaan vasta datanhallinnan toteuttamisen yhteydessä.

Datanhallinnan kattavuus määrittää toteutetaanko datanhallintaa yhdessä organisaatiossa, useamman yrityksen välillä vai jopa yritysekosysteemin sisällä. Se missä laajuudessa datanhallintaa toteutetaan, vaikuttaa siihen millaisia hallintamekanismeja organisaatio tarvitsee. Kohteena oleva data kuvaa, minkälainen data on datanhallintamallin keskiössä. Datanhallinnan kohteena olevan datan määrittely on tärkeää, sillä datan tyyppi voi asettaa vaatimuksia datanhallinnan suunnittelulle. Ero tulee tehdä erityisesti perinteisen datan hallinnan ja big datan hallinnan välillä. (Abraham et al. 2019) Perinteinen data muodostuu master datasta, metadatasta, transaktionaalisesta datasta ja referenssidatasta (Talbourt & Zhou 2015). Se rakentaa perustan organisaation ja sen tietojärjestelmien toiminnalle. Big datalle taas on tyypillistä sen moninaisuus, nopeus ja volyyymi (Lee et al. 2014). Se tarjoaa innovatiivisia mahdollisuuksia kehittää organisaation toimintoja tai uusia liiketoimintastrategioita tarjoamalla sellaista uutta ymmärrystä, jota perinteinen data ei voi tarjota (Lee et al. 2014). Kun perinteisen datan hallinta pyrkii usein varmistamaan datan johdonmukaisen käyttämisen koko organisaatiossa, big datan hallinta painottuu big dataan liittyvien haasteiden ratkaisuun, kuten yksityisyyden varmistamiseen ja datan epäjohdonmukaisuuksien ratkaisemiseen (Kim & Cho 2018). Näin ollen datanhallinnan kohteen määrittely on tärkeää, sillä se vaikuttaa datanhallintamallin painopisteisiin. Jos datanhallinnan kohdetta ei määritellä, malli voidaan muodostaa väärään tarkoitukseen sopivaksi, jolloin datanhallinnalla ei välttämättä saada haluttuja vaikutuksia.

Viimeisenä Abrahamin ja muiden (2019) mallissa ovat päätöksentekoalueet. Ne ovat alueita, joiden ympärillä organisaatio tekee päätöksiä datan hallintaan tai sen käsittelyyn liittyen ja joihin hallintamekanismeja sovelletaan (Khtari & Brown 2010; Alhassan et al. 2018; Abraham et al. 2019). Päätöksentekoalueiden määrittäminen ei tarkoita koko alueen vastuun määrittämistä tietylle yksikölle tai tietylle roolille. Khatrin ja Brownin (2010) mukaan päätöksenteko alueilla voi olla joko hajautettua tai keskitettyä eli vastuu tietyistä alueista voidaan keskittää yhdelle ryhmälle tai hajauttaa useille rooleille eri yksiköissä.

Päätöksentekoaalueet ovat datanhallinnassa tärkeässä osassa sillä monien organisaatioiden datanhallinnan tavoitteet liittyvät usein kahteen tai kolmeen päätöksentekoaalueeseen (Abraham et al. 2019).

Tässä työssä keskitytään edellä esiteltyistä datanhallinnan osa-alueista erityisesti hallintamekanismeihin sekä päätöksentekoaalueisiin, joilla niitä sovelletaan. Seuraavissa kappaleissa tutustutaan näihin tarkemmin. Päätöksentekoaalueiden suuren määrän vuoksi niiden syvällisempi tarkastelu on rajattu kohdeorganisaation kannalta keskeisimpien osa-alueiden tarkasteluun.

3.2.1 Menettelylliset hallintamekanismit

Datanhallinnan menettelyllisten hallintamekanismien tavoitteena on varmistaa, että data tallennetaan tarkasti, säilytetään turvallisesti, käytetään tehokkaasti ja jaetaan asianmukaisesti (Abraham et al. 2019). Khatrin ja Brownin (2010) mukaan menettelylliset hallintamekanismit, tai heidän mukaansa dataperiaatteet, varmistavat datanhallinnan ja liiketoiminnan välisen yhteyden sekä lainsäädännön ja säädösten noudattamisen. Tutkijoiden mukaan menettelylliset hallintamekanismit määrittävät toivotun käyttäytymisen niin IT-ammattilaisille kuin liiketoimintaosaajille. Yksi merkittävimmistä menettelyllisistä hallintamekanismeista on datastrategia.

Datastrategia on suunnitelma, joka määrittelee, miten organisaatio hallitsee, käyttää, analysoi, jakaa ja suojaa dataa saavuttaakseen strategiset liiketoimintatavoitteet (Chakravorty 2020; Medeiros et al. 2020). Datastrategia on merkittävä osa datanhallintaa, sillä Medeiros ja muut (2020) ovat todenneet positiivisen yhteyden datastrategian ja yrityksen kilpailuedun välillä. Heidän mukaansa datastrategian käyttöönotolla on merkittäviä vaikutuksia liiketoimintastrategioiden edistämiseen ja ohjaamiseen sekä yrityksen kilpailukyvyille. Vaikka datastrategian linjaaminen organisaation strategian ja tavoitteiden mukaan on tärkeää, Capgeminin toteuttaman tutkimuksen mukaan vain 41 % liiketoimintajohtajista energia-alalla kokee, että heidän yrityksensä data- ja analytiikkastrategia on linjassa yrityksen liiketoimintastrategian kanssa (Gapgemini Research Institute 2020).

Organisaatioiden datastrategia pyrkii usein tasapainottelemaan kahden asian välillä, datan suojaaminen ja datan hyödyntäminen (Dallemule & Davenport 2017; Medeiros et al. 2020). Dallemulen ja Davenportin (2017) mukaan se, kumpaa yrityksen datastrategia painottaa ja missä suhteessa, määräytyy yrityksen strategian, toimialan, kilpailullisen ympäristön ja sääntelyn mukaan. Kirjoittajien mukaan datan suojaamista painotetaan usein, jos yritykselle on tärkeää vähentää yleisiä toimintakuluja, noudattaa toimialan sääntelyvaatimuksia, estää tietomurtoja, pienentää operatiivisia riskejä ja parantaa datan laatua. Tämä vaatii usein yhtä totuutta tukevan arkkitehtuurin, vahvan datanhallinnan

ja keskitetyn päätöksenteon. Data hyödyntämistä taas painotetaan datastrategiassa usein, jos yritykselle on tärkeää muun muassa liikevaihdon kasvattaminen, uusien tuotteiden ja palveluiden luominen, nopea reagointi markkinamuutoksiin, edistyneen analytiikan hyödyntäminen liiketoiminnan parantamiseksi sekä uusien datalähteiden hyödyntäminen. Tällöin organisaatioille tyypillistä on usein korkea datan joustavuus eli kyky soveltaa dataa ja mukauttaa dataan liittyviä prosesseja tarpeiden mukaan. Lisäksi datan hyödyntämistä korostavalle datastrategialle on tyypillistä hajautettu päätöksenteko. Tällaista datastrategiaa toteuttavat organisaatiot saattavat myös hyödyntää muitakin kuin yhtä totuutta tukevia arkkitehtuureja. Dallemulen ja Davenportin mukaan yritysten asemoituminen datan suojaamisen ja hyödyntämisen suhteen pysyy kuitenkin harvoin vakiona. Tämä tukee aiemmin esitettyä käsitystä siitä, että datanhallinta ei ole kertaluontoisesti toteutettava projekti, vaan jatkuva prosessi, jota tulisi myös kehittää jatkuvasti.

Datastrategian lisäksi menettelyllisiin hallintamekanismeihin kuuluu muun muassa datapolitiikat, -standardit, -prosessit sekä menettelytavat. Datapolitiikat ovat lyhyitä linjauksia, jotka tarjoavat ylätasoa ohjeita ja sääntöjä datan luomiseen, yhdistämiseen, varastointiin, laatuun sekä käyttöön (Abraham et al. 2019; Alhassan et al. 2019). Datapolitiikat tukevat datastandardeja. Kun datapolitiikat kuvaavat mitä pitää tai ei pidä tehdä, datastandardit kuvaavat miten tehdään. (DAMA 2017) Datastandardit varmistavat, että datan esittäminen ja dataan liittyvien toimintojen toteuttaminen on yhdenmukaista ja normalisoitua koko organisaation tasolla (Abraham et al. 2019). Datastandardien lisäksi myös dataprosessit ja menettelytavat ovat yhteydessä datapolitiikkoihin. Datapolitiikat tulisi määrittellä tarkemmin ja tuoda toteutukseen dataprosesseihin ja menettelytapoihin liittyvien toimintojen kautta (Alhassan et al. 2019). Dataprosessit ovat standardisoituja, dokumentoituja ja toistettavissa olevia menetelmiä, joita käytetään datan hallintaan. Dataan liittyvät menettelytavat taas ovat dokumentoituja metodeja, tekniikoita ja vaiheita, joita noudatetaan tietyn tehtävän suorittamisessa. (Abraham et al. 2019).

3.2.2 Rakenteelliset hallintamekanismit

Rakenteelliset hallintamekanismit määrittelevät raportointirakenteet, hallintoelimet ja vastualueet. Ne kattavat roolien ja vastuiden määrittämisen sekä päätöksenteon allokoinnin. Aiemmin datanhallinta on ollut pitkälti IT-osastojen ja teknologioiden vastuulla, mutta datamäärien kasvaessa sen hallinta on vaikeutunut. Enää datanhallintaa ei voida jättää vain IT-osaston vastuulle, vaan tarvitaan kokonaisvaltaisempia lähestymistapoja. (Al-Ruithe et al. 2019; Benfeldt et al. 2020) DAMA:n (2017) mukaan datanhallinta on tehokkainta, kun se on koko organisaation kattava toiminto, eikä vain yhden yksikön vastuulla. Eri datanhallinnan vastuita voidaan silti toteuttaa joko keskitetysti tai hajautetusti

(Khatri & Brown 2010). Tilivelvollisuudet ja vastuut täytyy kuitenkin olla selkeästi määritelty. Selkeät vastuut ja roolit on tunnistettu useissa tutkimuksissa yhdeksi datanhallinnan onnistumisen kannalta kriittiseksi tekijäksi (Mahanti 2018; Al-Ruithe et al. 2019; Alhassan et al. 2019). Taulukossa 5 on listattu datanhallintaan yleisimmin liitettäviä rooleja ja niiden kuvaukset. Rooleja tarkasteltaessa tulee huomioida, että ne ovat organisaatiokohtaisia ja jokaisen organisaation tulee määrittää itselleen sopivat roolit (Mahanti 2018).

Taulukko 5. Yleiset datanhallinnan roolit

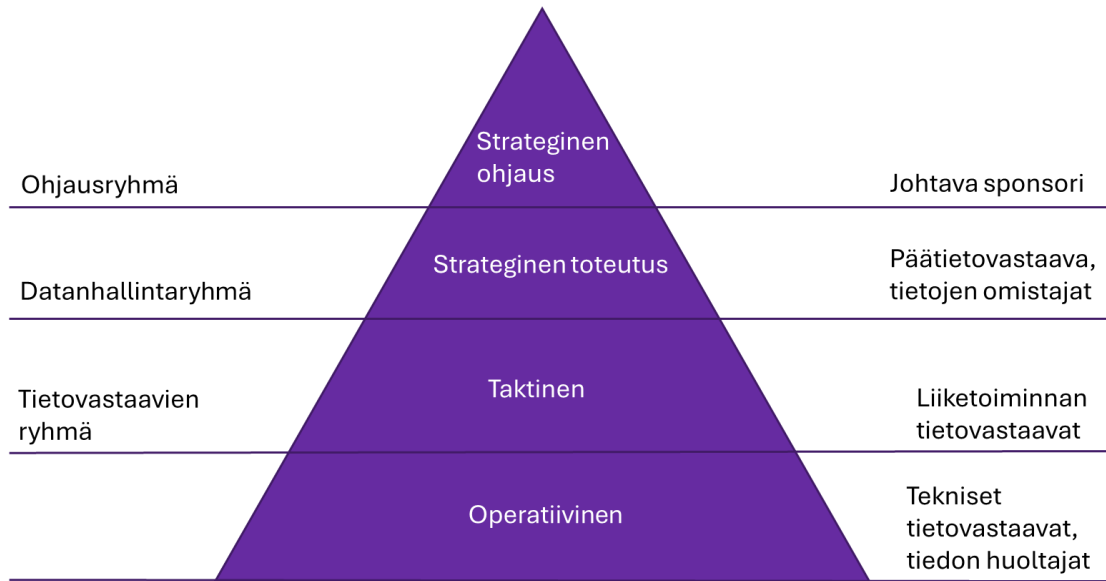
Rooli	Kuvaus	Lähde
Johtava sponsori (engl. executive sponsor)	Johtava sponsori tarjoaa strategisen tason näkemystä, priorisointia liiketoiminnan kannalta sekä rahoitusta datanhallinnan projekteille. Hän on ideaalissa tapauksessa yksi yrityksen korkeimman tason johtajista.	Weber et al. 2009, Abraham et al. (2019)
Päädatavastaava (engl. chief steward)	Päädatavastaava johtaa datanhallintaryhmää. Hänen vastuullaan on laittaa datanhallintaryhmässä tehdyt päätökset käytäntöön.	Korhonen et al. (2013)
Datan omistaja (engl. Data owner)	Datan omistaja on vastuussa liiketoimintayksikkönsä dataomaisuuksista. Hän kommunikoi datan vaatimuksista ja riskeistä ja on vastuussa omistamansa datan laadusta.	Khatri & Brown (2010), Mahanti (2018), Abraham et al. (2019)
Liiketoiminnan tietovastaava (engl. Business data steward)	Liiketoiminnan tietovastaavilla on tietämystä tietystä liiketoiminta-alueesta ja datan vaatimuksista. He osaavat muuttaa nämä vaatimukset teknisiksi vaatimuksmäärittelyiksi.	Khatri & Brown (2010), DAMA (2017), Mahanti (2018), Abraham et al. (2019)
Tekninen tietovastaava (engl. Technical data steward)	Tekniset tietovastaavat ovat tietotekniikan ammattilaisia ja he toimivat liiketoiminnan tietovastaavien vastapuolina.	DAMA (2017), Abraham et al. (2019)
Tiedon huoltaja (engl. Data custodian)	Tiedon huoltaja on vastuussa datan ylläpitämisestä tietojärjestelmissä linjassa liiketoiminnan vaatimusten kanssa.	Khatri & Brown (2010), Mahanti (2018)

Osa taulukossa 5 esitellyistä rooleista voi kuitenkin olla päällekkäisiä ja yhdelle ihmiselle voi kuulua useampi rooli. Esimerkiksi datan omistajat voidaan tulkita liiketoiminnan tietovastaaviksi, joilla on lisäksi valta tehdä päätöksiä oman alueensa datasta. (DAMA 2017) Datanhallinnan rooleihin liitetään usein eri organisaation toimintoja yhdistäviä ryhmiä, jotta voidaan varmistaa yhtenäisten linjausten noudattaminen koko organisaatiotasolla. Mahantin (2018) mukaan tämä on tärkeää, koska datanhallinta koskettaa myös dataa, jota käytetään useissa liiketoimintayksiköissä. Tutkijan mukaan koko organisaation kattavien hallintaelimien muodostaminen varmistaa, että yksiköt tekevät yhteistyötä ja dataa käsitellään tehokkaasti ja sen laatu varmistetaan. Taulukossa 6 on esitelty näistä ryhmistä yleisimmät.

Taulukko 6. Yleiset organisaation toimintoja yhdistävät datanhallinnan ryhmät

Rooli	Kuvaus	Lähde
Ohjausryhmä (engl. steering committee)	Ohjausryhmä vastaa organisaation strategisen suunnitelman ja politiikan hyväksymisestä ja rahoittamisesta.	DAMA (2017), Mahanti (2018)
Datanhallintaryhmä (engl. DG council)	Datanhallintaryhmä on hierarkioita ylittävä organisaation eri toimintoja yhdistävä hallintoelin. Se asettaa strategisen suunnan organisaation koko datanhallinnalle ja varmistaa, että se on linjassa organisaation tavoitteiden kanssa. Ryhmä koostuu usein päätietovastaavasta, datan omistajista sekä mahdollisesti myös tietovastaavista.	Khatri & Brown (2010), Mahanti (2018), Abraham et al. (2019), Chakravorty (2020)
Tietovastaavien ryhmä (engl. Data steward council)	Ryhmän tarkoitus on tuoda tietovastaavat yhteen tekemään dataan liittyviä päätöksiä esimerkiksi datan määrittelyistä ja standardeista projekteihin liittyen. Ryhmä tukee ja valvoo datanhallintaryhmän ja päädatavastaavan käynnistämiä aloitteita.	Korhonen et al. (2013), Mahanti (2018),

Datanhallinnan roolit organisaatiossa kattavat tyypillisesti kolme tasoa: strateginen, taktinen ja operatiivinen (Korhonen et al. 2013; DAMA 2017; Chakravorty 2020). Näistä strateginen taso voidaan erottaa vielä strategisen ohjauksen ja toteutuksen tasoon. Strategisessa ohjauksessa suunnitteluhorisontti on yli viisi vuotta. Tällä tasolla päätökset liittyvät organisaation liiketoimintamalleihin, pitkän aikavälin tavoitteisiin, tuleviin suuntiin ja niissä täytyy huomioida usein myös ulkoisten tekijöiden vaikutus. Strategisessa toteutuksessa suunnitteluhorisontti on 2–5 vuotta. Strategiset aikomukset jalostetaan konkreettisemmiksi tavoitteiksi ja suunnitelmiksi operatiivisille yksiköille. Strategia optimoidaan integroimalla useita toimintoja ja eri toimintoja yhdistäviä prosesseja. Tämä saavutetaan organisaatiotason suunnitelmien ja järjestelmien avulla. Taktisen tason suunnitteluhorisontti sen sijaan on yhdestä kahteen vuoteen. Tämän tason päätöksentekooikeudet rajoittuvat usein ydinliiketoimintaprosessien tehokkuuden parantamiseen. Tasolla yhdistetään tiimejä eri toiminnoista kehittämään olemassa olevia työtapoja ja prosesseja. Operatiivisen tason suunnitteluhorisontti on esitellyistä tasoista lyhyin, kolmesta kuukaudesta vuoteen. Operatiivinen päätöksenteko liittyy välittömän tulevaisuuden haasteisiin, sillä on suora vaikutus liiketoimintaan ja se keskittyy usein jatkuvaan kehittämiseen ja laatuun. (Korhonen et al. 2013) Kuvassa 6 on havainnollistettu yleisiä datanhallinnan rooleja ja niiden strategiset tasot.



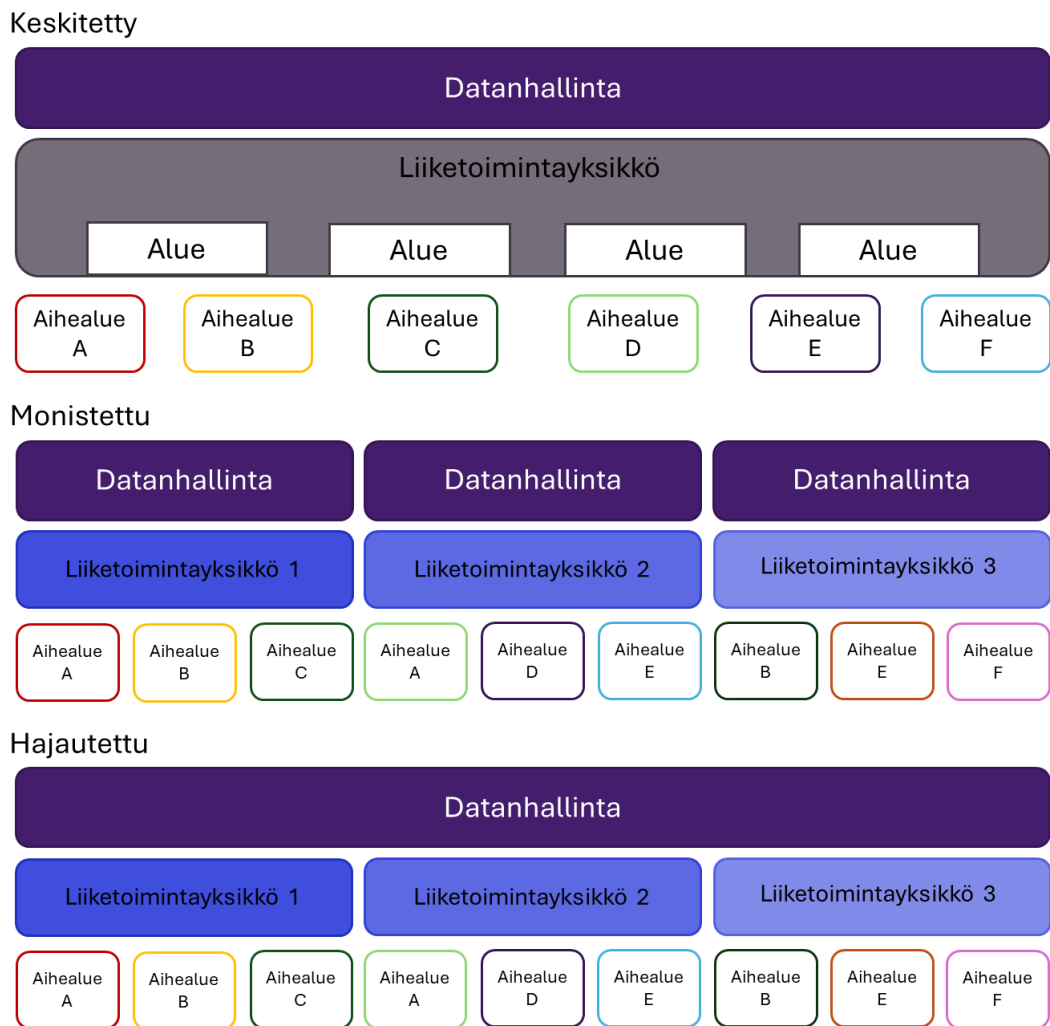
Kuva 6. *Datanhallinnan yleisiä rooleja ja niiden strateginen taso (mukailen Korhonen et al. 2013)*

Näiden strategisten tasojen huomioiminen on tärkeää, jotta datanhallinnan yhteys yrityksen strategiaan ja tavoitteisiin voidaan varmistaa. Jos datanhallintaa toteutetaan vain operatiivisella tasolla, tämä yhteys jää todennäköisesti toteutumatta. Tällöin toiminta ohjautuu helposti välittömien ongelmien ratkaisemiseen, eikä toimintaa pystytä suunnittelemaan pitkällä aikavälillä. Kaikkien tasojen mukaan ottaminen varmistaa, että yrityksen strategiset tavoitteet siirtyvät myös datan hyödyntämiseen liittyvään päivittäiseen toimintaan.

Kuten aiemmin todettu, rakenteelliset hallintamekanismit kattavat roolien ja vastuiden määrittämisen lisäksi päätöksenteko-oikeuksien allokoinnin. Päätöksenteko-oikeuksien allokointi määrittää, millä organisaation yksiköllä on valta ja vastuu hallita datanhallinnan toimintoja. Abraham ja muut (2019) tunnistavat datanhallinnan päätöksenteko-oikeuksissa hierarkkisen asemoitumisen, toiminnallisen asemoitumisen sekä päätöksenteon hajautuneisuuden tason. Tutkijoiden mukaan hierarkkinen asemoituminen määrittää, millä organisaation hierarkiatasolla päätökset tehdään. Toiminnallinen asemointi määrittää, mikä osasto vastaa päätöksenteosta. Päätöksenteon hajautuneisuus taas määrittää, tehdäänkö päätökset keskitetysti yhdessä toiminnossa vai hajautetusti eri toiminoissa.

DAMA (2017) on tunnistanut kolme erilaista mallia datanhallinnan hajautukseen liittyen: keskitetty (engl. centralized), monistettu (engl. replicated) ja hajautettu (engl. federated). Jo aiemmin Weber ja muut (2009) ovat tunnistaneet näistä keskitetyn ja hajautetun mallin. DAMA:n (2017) mukaan keskitetyssä mallissa yksi datanhallinnan organisaatio val-

voo kaikkia toimintoja kaikilla aihealueilla. Monistetussa mallissa kaikissa liiketoimintayksiköissä on käytössä yksi sama datanhallintamalli ja standardit, mutta eri yksiköitä yhdistävää datanhallinnan organisaatiota ei ole. Sen sijaan hajautetussa mallissa yksi datanhallinnan organisaatio koordinoi useiden liiketoimintayksiköiden kanssa, jotta organisaatiossa voidaan ylläpitää yhdenmukaiset määritelmät ja standardit. Nämä mallit on havainnollistettu kuvassa 7.



Kuva 7. Datanhallinnan hajautus (mukaillen DAMA 2017)

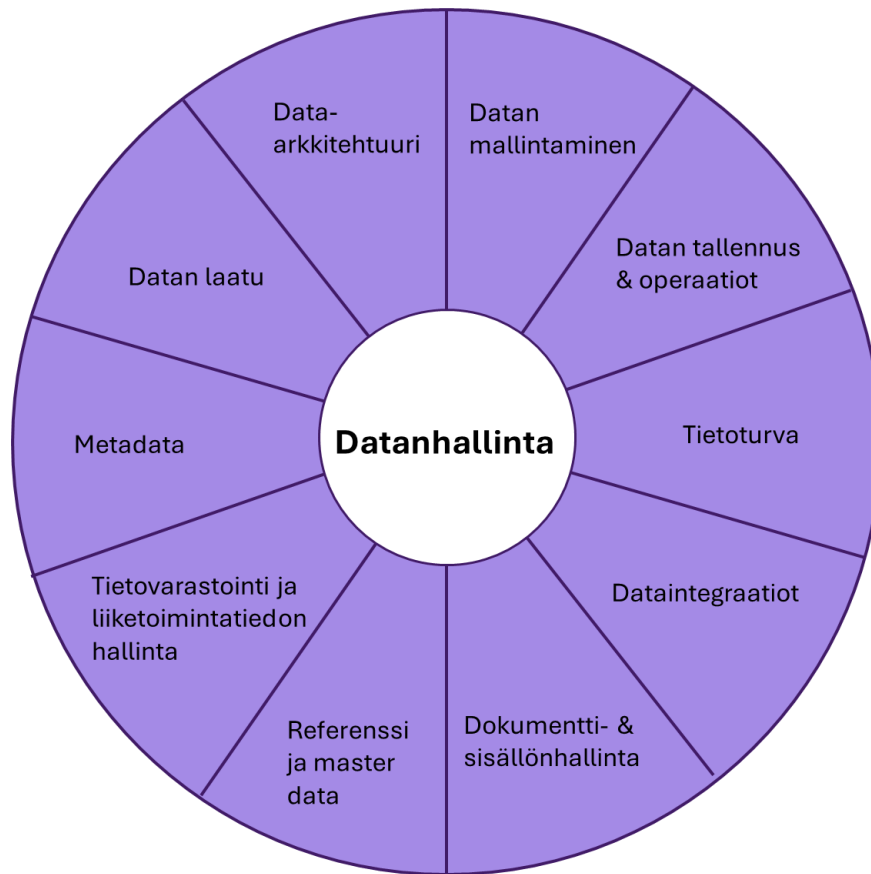
Yleisesti ottaen keskitetyn hallintorakenteen ja päätöksenteon eduksi voidaan katsoa tehokkuus, parempi kontrolli, päätöksenteon nopeus ja selkeä vastuunjako. Sen sijaan hajautetun hallintorakenteen ja päätöksenteon eduksi voidaan nähdä alempien organisaatiotasojen motivaation ja sitoutumisen paraneminen, innovatiivisuuden lisääntyminen sekä joustavuus ja parempi kyky toimia monimutkaisissa ja muuttuvissa ympäristöissä. (Altamimi et al. 2023) Näitä samoja etuja voidaan tunnistaa myös datanhallinnan hajautukseen liittyen. Weberin ja muiden (2009) mukaan datanhallinnassa keskitetyn raken-

teen eduksi voidaan tunnistaa parempi kontrolli, kun taas hajautettu rakenne mahdollistaa paremman joustavuuden, liiketoimintatarpeisiin vastaamisen sekä jokaisen yksikön tarpeisiin mukautetut ratkaisut.

Datanhallinnan hajautusmallien erilaisten vahvuuksien perusteella voidaan todeta, että muodostaessaan ja kehittäessään datanhallintaa yritysten tulee arvioida organisaation tavoitteita ja arvioida, minkälaiset roolit ja päätöksenteon hajautusmalli palvelee niiden saavuttamista. Jotkin yritykset tarvitsevat erityisesti parempaa kontrollia, jolloin keskiteympi datanhallintamalli voi sopia heidän tarpeisiinsa paremmin (Altamimi et al. 2023). Jos taas yrityksen tavoitteiden täytyminen vaatii innovointikyvykkyyttä ja nopeaa reagoitua toimintaympäristön muutoksiin, hajautetumpi datanhallinnan rakenne voi palvella yritystä paremmin (Altamimi et al. 2023). Datanhallinnan hajautuksen aste määrittää myös osin, mitä rooleja yritys tarvitsee. DAMA:n (2017) ehdottamassa monistetussa datanhallintamallissa koko organisaation kattaville hallintoelimille ei ole tarvetta, kun taas hajautetussa mallissa ne varmistavat, että datanhallintaa voidaan koordinoida tehokkaasti eri liiketoimintayksiköiden kanssa.

3.2.3 Päätöksentekoalueet

Datanhallinnan päätöksentekoalueiksi voidaan tulkita datan operatiivisen hallinnan toiminnot, koska kuten luvussa 3.1 todettiin, datanhallinta voidaan käsittää niitä ohjaavana organisoitumismallina. Datanhallinta määrittää rakenteen ja säännöt, joiden avulla datan operatiivisen hallinnan toiminnot toteutetaan tehokkaasti ja säännöstenmukaisesti. DAMA:n (2017) mukaan datanhallinta ohjaa kymmentä datan operatiivisen hallinnan toimintoa, joita ovat datan mallintaminen, datan tallennus ja operaatiot, tietoturva, dataintegraatiot, dokumentti- ja sisällönhallinta, referenssi ja master data, tietovarastointi ja liiketoimintatiedonhallinta, metadata, datan laatu ja data-arkkitehtuuri. Nämä alueet on esitelty kuvassa 8. Khatri ja Brown (2010) sen sijaan ehdottavat datanhallinnan päätöksentekoalueiksi datan laatua, metadataa, datan käyttöä ja datan elinkaarta, joita yhdistävät luvussa 3.2.1 mainitut dataperiaatteet. Tunnistetuissa päätöksentekoalueissa on siis eroja, mutta myös yhtäläisyyksiä. Molemmat tunnistavat datan laadun, metadatan, ja datan käytön ja turvallisuuden merkityksen. Lisäksi datan elinkaariajattelu liittyy moneen datan operatiivisen hallinnan toimintoon, kuten esimerkiksi dokumentti- ja sisällönhallintaan ja datan tallennukseen ja operaatioihin.



Kuva 8. Datanhallinnan suhde datan operatiiviseen hallintaan (mukaan DAMA 2017)

Koska kirjallisuudesta voidaan tunnistaa useita datanhallinnan päätöksentekoaalueita, työssä tarkastellaan niistä tarkemmin vain kohdeorganisaation datanhallintamallin kannalta oleellisimpia. Oleellisimpien osa-alueiden tunnistus tehtiin kohdeorganisaation kanssa käytyjen keskusteluiden sekä jo aiemmin esitellyn strategiadokumentin perusteella. Tarkemman tarkastelun kohteeksi valittiin datan laatu, datan mallintaminen sekä tietovarastointi ja liiketoimintatiedon hallinta. Nämä kaikki ovat yhteydessä kohdeorganisaation data- ja analytiikkakehitykseen, jolla on tiivis yhteys yhtiön datanhallintaan.

Datan laatu tarkoittaa datan kykyä täyttää sen käytölle asetetut vaatimukset tietyssä kontekstissa. Datan laatua voidaan kuvata erilaisten ulottuvuuksien kautta, joita ovat muun muassa datan tarkkuus, hyödynnettävyys, saatavuus ja ajantasaisuus. (Mahanti 2019) Datan laatu -pätöksentekoaalue sisältää muun muassa standardien asettamisen datan laadun ulottuvuuksille. Lisäksi alueeseen voidaan liittää datan laadun arvioinnin mahdollistavien prosessien luominen. (Khatri & Brown 2010; Abraham et al. 2019) Datan laatua arvioitaessa tulee kuitenkin muistaa, että datan laadun taso on usein sidottu suoritettavan tehtävän kontekstiin (Mahanti 2019). Joissakin tapauksissa samana päivänä päivitetty data on ajantasaisuudeltaan riittävä, mutta toisissa se voi olla jo vanhentunutta tietoa.

Datan mallintaminen voidaan tulkita kokoelmaksi konseptuaalisia työkaluja, joilla voidaan kuvata dataa, sen suhteita, semantiikkaa ja rajoituksia. Datan mallinnuksessa voidaan tunnistaa kolme tasoa: käsitelmallinnus, looginen mallinnus ja fyysinen mallinnus. Käsitelmallinnus tarkoittaa entiteettien ja niiden suhteiden tunnistamista ja kuvaamista sekä havaintojen merkitsemistä esimerkiksi entiteetti-suhdekaavioon. Looginen mallinnus taas tarkoittaa tietokannan taulujen määrittämistä taustalla olevan relaatiomallin mukaan. Viimeisenä fyysinen mallinnus käsittää tietokannan optimoinnin. (Schreibman et al. 2016)

Tietovarastoinnin ja liiketoimintatiedon hallinnan yksi tavoite on tukea tarkkaa ja ajantasaista päätöksentekoa. Tietovarastointi käsittelee sitä, miten useista lähteistä voidaan yhdistää ja varastoida suuret määrät monimuotoista dataa yhteen paikkaan. Data varastoidaan muodossa, jossa sitä voidaan helposti hakea ja hyödyntää. Tietovarastoinnin yksi merkittävimmistä tehtävistä on liiketoimintatiedon hallinnan tukeminen. Liiketoimintatiedon hallinta on kokoelma teknologioita ja menetelmiä, jotka auttavat yrityksiä ymmärtämään liiketoimintaa, markkinoita ja tekemään oikeita päätöksiä. Liiketoimintatiedon hallinnassa hyödynnetään usein tietovarastoa tietolähteenä ja pyritään esittämään dataa ja tekemään analyyskejä päätöksenteon tueksi. (Gaol et al. 2023)

Datanhallinnan päätöksentekoalueet auttavat määrittämään mitä dataan liittyviä päätöksiä organisaatiossa tulisi tehdä, jotta voidaan mahdollistaa arvontuottaminen datasta. Niitä voidaan käyttää vastuiden jakamisen apuna ja näin vähentää riskiä sekaannuksille ja olennaisten osa-alueiden laiminlyönnille. Esitellyt päätöksentekoalueet ovat kattavia, mutta kuten Abrahamin ja muiden (2023) tutkimuksessa on tehty, päätöksentekoalueita tulisi tarkastella ja soveltaa kontekstikohtaisesti.

3.3 Datanhallintamallin muodostamisessa huomioitavat taustatekijät

Datanhallinta on jatkuva prosessi, jonka täytyy pystyä sopeutumaan sisäisiin ja ulkoisiin muutoksiin. Sen täytyy mukautua uusiin sisäisiin datatarpeisiin sekä muutoksiin priorisoinnissa, liiketoimintaympäristössä ja ulkoisissa vaatimuksissa, kuten laissa ja säännöksissä. (DAMA 2017; Abraham et al. 2019; Mahanti 2019) Tarvetta datanhallinnan mukautuvuudelle tukee myös aikaisemmin esitetty näkemys siitä, että yritysten asemoituminen datan suojaamisen ja hyödyntämisen suhteen pysyy harvoin vakiona. Jos yrityksen datastrategia muuttuu, täytyy koko datanhallintaa arvioida silloin uudestaan.

Kirjallisuudessa on tunnistettu tekijöitä, jotka vaikuttavat datanhallinnan suunnitteluun, käyttöönottoon ja omaksumiseen. Näitä ovat yrityksen datanhallinnan tavoitteet, jotka johdetaan yrityksen liiketoimintastrategiasta ja tavoitteista (DAMA 2017; Abraham et al.

2019; Chakravorty 2020; Lis et al. 2023) sekä ulkoiset ja sisäiset kontekstitekijät (Weber et al. 2009; Otto 2011b; Abraham et al. 2019; Lis et al. 2023). Weberin ja muiden (2009) mukaan datanhallinta voi onnistua vain silloin, kun datanhallintamalli on sopiva organisaatiokohtaisten kontekstitekijöiden kanssa. Tämä seurauksena voidaan olettaa, että myös kontekstitekijöissä tapahtuvien muutosten jälkeen datanhallintamallin sopivuutta tulisi arvioida uudelleen.

3.3.1 Datanhallinnan tavoitteet

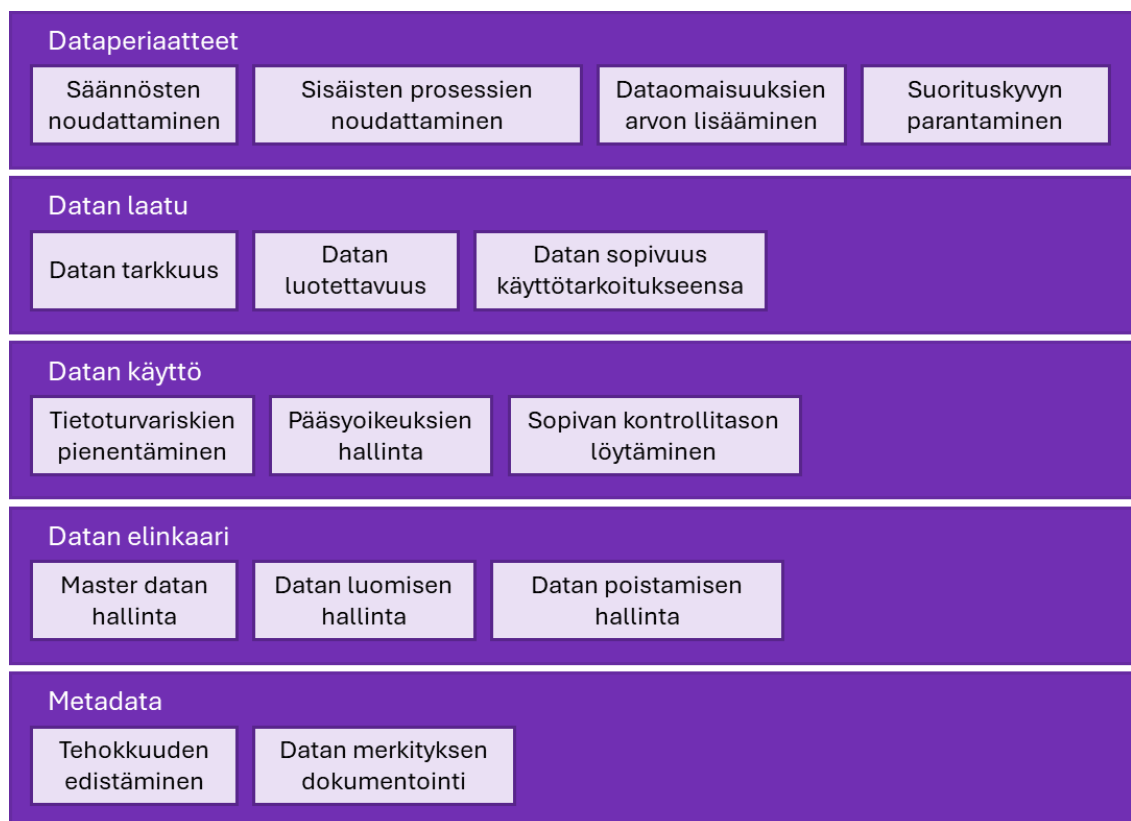
Organisaation datanhallinta tulisi suunnitella tukemaan liiketoimintastrategiaa ja -tavoitteita (DAMA 2017; Abraham et al. 2019; Chakravorty 2020; Lis et al. 2023). Tällöin datanhallinnan tavoitteidenkaan ei tulisi olla irrallisia organisaation muusta toiminnasta vaan niiden tulisi perustua organisaation strategiaan ja liiketoimintatavoitteisiin (Otto 2011b; Chakravorty 2020). Jos organisaation tavoitteena on parantaa asiakastyytyväisyyttä, sen täytyy pystyä hyödyntämään asiakasdataa tehokkaasti eli organisaation asiakasdatan täytyy olla hyvälaatuista. Tällöin datanhallinnan tavoitteena on todennäköisesti asiakasdatan laadun parantaminen.

Tämä johtaa siihen, että organisaatiot eivät voi hyödyntää yhtä samaa datanhallintamallia, koska niillä on erilaiset liiketoimintastrategiat ja näin ollen myös erilaiset tavoitteet datanhallinnalle. Joidenkin yritysten datanhallinnan tarkoitus voi olla datan laadun parantaminen, kun taas toiset voivat nostaa datanhallinnan keskiöön tietosuojan. Nämä tavoitteet vaikuttavat siihen, minkälaisen datanhallintamallin yritykset muodostavat. Sensitiivisen asiakasdatan turvaaminen voi vaatia virallisen ja tarkasti määritellyn datanhallintamallin, kun taas datan laadun parantamiseen tähtäävät yritykset saattavat hyödyntää jotain kevyempää ja taktisempaa. (Jimenez et al. 2019) Datanhallintamallien sisältö vaihtelee merkittävästi sen mukaan, mikä on niiden fokuksessa (Mahanti 2018).

Siitä huolimatta, että organisaatioiden datanhallinnan tavoitteet vaihtelevat, DAMA (2017) on tunnistanut yleisiksi datanhallinnan tavoitteiksi riskien vähentäminen tai prosessien parantaminen. DAMA:n esittämässä yleisissä datanhallinnan motivaattoreissa ja tavoitteissa korostuu tietoturvan, säännösten noudattamisen, datan laadun, metadatan hallinnan ja dataan liittyvien sopimusten hallinnan parantaminen. Walsh ja muut (2022) ovat tutkimuksessaan tunnistaneet osittain samoja tekijöitä. He ovat luokitelleet yleiset datanhallinnan motivaatiot ja tavoitteet Khattrin ja Brownin (2010) päätöksentekoa-alueiden mukaan, joista tutkimuksen tuloksissa korostuu dataperiaatteiden, datan laadun ja datan käytön merkitys. Walshin ja muiden (2022) mukaan organisaatioiden datanhallinnan tavoitteet dataperiaatteisiin liittyen koskevat pääasiassa säännösten sekä sisäisten liike-

toiminta- ja dataprosessien noudattamisen parantamista. Sisäisten prosessien noudattamisen lisäämisellä tavoitellaan päätöksenteon parantamista, mikä tehostaa liiketoimintaa ja johtaa lopulta organisaation tuloksen kasvuun.

Keskeinen osa tehokasta päätöksentekoa on tarkka data, johon päätöksentekijät voivat luottaa. Walshin ja muiden (2022) tutkimuksessa tunnistetut keskeisimmät datan laatuun liittyvät datanhallinnan tavoitteet ja motivaatiot koskevatkin pääasiassa datan tarkkuuden ja luotettavuuden parantamista. Kolmannen keskeisen päätöksentekoalueen, datan käytön, motivaatiot liittyvät pääasiassa tietoturvariskien pienentämiseen ja datan käyttöoikeuksien hallintaan. Tutkimuksen mukaan metadatan ja datan elinkaaren rooli on datanhallinnan motivaattorina selkeästi pienempi kuin edellä mainittujen. Metadatan ja datan elinkaaren hallinnan pieni rooli motivaattoreissa on Walshin ja muiden (2022) mukaan yllättävää, koska ne ovat olennaisia osa-alueita datan laadunhallinnan kannalta, joka kuitenkin tunnistettiin yhdeksi pääasiallisista motivaattoreista datanhallinnalle. Walshin ja muiden tunnistamat keskeisimmät datanhallinnan motivaattorit ja tavoitteet kaikille päätöksentekoalueille on esitetty kuvassa 9.



Kuva 9. Organisaatioiden keskeiset motivaatiot datanhallinnan käyttöönottoon ja toteuttamiseen (mukaillen Walsh et al. 2022)

Vaikka datanhallinnan tavoitteiden tulisi perustua strategiaan ja liiketoimintatavoitteisiin, voivat ne todellisuudessa olla joskus osin ristiriidassa keskenään. Liiketoimintatavoitteena voi olla tuloksen kasvattaminen, kun taas datanhallinnan tavoitteena voi olla datan parempi turvaaminen, joka voi vaatia kalliiden toimenpiteiden toteuttamista. Tällaisia ristiriitoja voi syntyä myös tiimien tai yksiköiden erilaisista tavoitteista. Tällöin organisaation täytyy pystyä arvioimaan, miten erilaiset datanhallinnan ratkaisut auttavat strategisten liiketoimintatavoitteiden saavuttamisessa, miksi niitä tarvitaan ja miten ristiriidassa olevia tavoitteita voidaan priorisoida. (Sothilingam 2021 et al.)

3.3.2 Datanhallinnan muut kontekstitekijät

Tavoitteiden lisäksi organisaatioiden tulisi ottaa datanhallinnan muodostamisessa huomioon muut kontekstitekijät, jotka vaikuttavat datanhallinnan suunnitteluun, käyttöön ottoon ja omaksumiseen. Kontekstitekijöiden huomioiminen mahdollistaa organisaatioille datanhallinnan muokkaamisen omiin ympäristöihin ja tarpeisiin sopiviksi. (Weber et al. 2009; Otto 2011b; Abraham et al. 2019; Lis et al. 2023) Kirjallisuudessa nämä kontekstitekijät on jaettu ulkoisiin ja sisäisiin tekijöihin. Merkittävimmät ulkoiset tekijät kattavat lait ja sääntelyn sekä yrityksen toimialan (Abraham et al. 2019; Lis et al. 2023). Sisäiset kontekstitekijät taas kattavat muun muassa organisaatorakenteen, kilpailustrategian, prosessien harmonisointiasteen, päätöksentekotyylin ja -kulttuurin sekä datan tyypin ja kasvuvauhdin (Weber et al. 2009; Lis et al. 2023). Organisaation koko datanhallinnan kontekstitekijänä on Weberin ja muiden (2009) mukaan kiistelty aihe, josta löytyy näkemyksiä sekä puolesta että vastaan. Tässä työssä ei löydetty tutkimuksia, jotka olisivat luokitelleet yrityksen koon varsinaiseksi kontekstitekijäksi. Datanhallinnan viitekehyksiä sovellettaessa tulee kuitenkin huomioida, että organisaation koolla voi olla vaikutusta datanhallinnan suunnitteluun. Esimerkiksi Devoksen ja muiden (2012) mukaan hallinta-teorioita, jotka selittävät ilmiötä suuressa yrityksessä ei voida usein soveltaa sellaiseen pienissä ja keskisuurissa yrityksissä. Pienillä ja keskisuurilla yrityksillä on usein käytössään vähemmän resursseja, mikä tekee niiden datanhallinnan toteuttamisesta haastavampaa (Begg & Caira 2012). Kirjallisuudesta tunnistetut datanhallinnan kontekstitekijät ja niiden kuvaukset on koottu taulukkoon 7.

Taulukko 7. *Datanhallinnan ulkoiset ja sisäiset kontekstitekijät (mukaillen Weber et al. 2009; Lis et al. 2023)*

Tekijä	Tekijän selitys
Lait ja sääntely	Miten datan käyttöä säädellään
Toimiala	Toimiala ja alueelliset ominaisuudet
Organisaation rakenne	Organisaation ja organisaation hallinnan hajautusaste
Kilpailustrategia	Suhtautuminen tuote- ja markkinakehitykseen ja sitoutuminen pysyvyyteen
Prosessien harmonisointiaste	Kuvaa, onko yrityksen prosessit yleisimmin määritelty ylätasolla vai määritelläkö prosessit paikallisesti

Päätöksentekotyylit ja -kulttuuri	Kuvaa, missä ja miten päätökset tehdään ja millainen organisaatiokulttuuri on datan arvostamisen suhteen
Datan tyyppi ja kasvuvauhti	Kuvaa datamäärien kasvuvauhdin sekä datatyyppien moninaisuuden

Kontekstitekijät vaikuttavat erityisesti datanhallinnan hajautuksen asteeseen. Esimerkiksi tiukasti säädellyt toimialat vaativat keskitetymmän hallintorakenteen kuin vähemmän säädellyt toimialat. Jos taas organisaatio keskittyy voiton tavoitteluun, se voi hyötyä enemmän keskitetystä datanhallinnasta, kun taas kasvua tavoittelevat yritykset voivat hyötyä hajautetusta datanhallinnasta. Sen sijaan yritykset, joilla on globaalisti harmonisoidut prosessit, voivat hyödyntää paremmin keskitettyä datahallintaa kuin yritykset, joilla on paikallisesti määritellyt prosessit. (Abraham et al. 2019) Jotta datanhallinta tukisi organisaation toimintaa, datanhallintamallissa tulisi huomioida edellä esitellyt kontekstitekijät ja mallin tulisi olla linjassa niiden kanssa. Jos esimerkiksi yrityksen toimiala on erittäin säädelty, mutta yrityksen datanhallintamalli ei tue datan käytön tiukkaa kontrollia, voivat löyhän hallinnan takia saatavat sanktiot kumota muut datanhallintamallilla saavutetut hyödyt.

3.4 Datanhallinnan toteuttaminen

Organisaatiokohtaisen datanhallintamallin muodostamisen tärkeyden lisäksi kirjallisuudessa on tunnistettu erilaisia datanhallinnan toteuttamiseen liittyviä kriittisiä menestystekijöitä. Kriittinen menestystekijä on termi, jota käytetään elementistä, joka on oleellinen organisaation tai projektin mission saavuttamisen kannalta. Datanhallinnan toteuttamisen kontekstissa kriittiset menestystekijät tarkoittavat keskeisiä elementtejä, joiden puuttuminen voi todennäköisesti johtaa datanhallinnan ja siihen liittyvien projektien epäonnistumiseen. (Mahanti 2018) Tällaisia kriittisiä menestystekijöitä ovat muun muassa johdon tuki (DAMA 2017; Mahanti 2018; Al-Ruithe et al. 2019), työntekijöiden sitoutuminen (Al-Ruithe et al. 2019; Karkošková 2023), henkilöstön kyvykkyydet (Mahanti 2018; Al-Ruithe et al. 2019) sekä kommunikaatio ja yhteistyö (DAMA 2017; Mahanti 2018; Bento et al. 2022). Tässä työssä näihin tekijöihin viitataan jatkossa datanhallinnan toteuttamisen edellytyksinä.

3.4.1 Johdon tuki

Datanhallinnalla on vaikutuksia yritysten yhteen keskeisimpään resurssiin, dataan. Tämän vuoksi datanhallintaan liittyvät kehitystoimenpiteet edellyttävät mandaattia koko organisaation laajuisille muutoksille. Tällaisia muutoksia ovat muun muassa muutokset työntekijöiden toimintatavoissa, asenteissa ja kulttuurissa. Näiden muutosten läpiviemiseksi tarvitaan johdon tuki ja hyväksyntä. Yrityksen ylimmän johdon tulee ymmärtää

datanhallinnan tärkeys, tukea siihen liittyviä aloitteita sekä edistää kulttuurillisia muutoksia ja yrityksen sitoutumista datanhallintaan. (Mahanti 2018; Chakravorty 2020)

Tsavatewa (2023) esittää, että jatkuvan tuen puute on yksi merkittävimmistä syistä, miksi datanhallinta epäonnistuu. Hänen mukaansa johdon jatkuva tuki on tärkeää varsinkin silloin, kun johtoon ja rooleihin tehdään muutoksia. Jatkuvan tuen lisäksi eri tason johdolla ja johtajilla tulee olla sama käsitys datanhallinnan tukemisesta sekä onnistuneen datanhallinnan määritelmästä. Jos johtajien välillä on suuria erimielisyyksiä, se voi lähettää muulle organisaatiolle ristiriitaisia viestejä ja lopulta johtaa datanhallinnan vastustamiseen ja datanhallinnan toteuttamisen epäonnistumiseen. (Mahanti 2018)

3.4.2 Työntekijöiden sitouttaminen datanhallintaan

Datanhallinta edellyttää koko henkilöstön osallistumista ja sitoutumista (Weber et al. 2009; Al-Ruithe et al. 2019). Tsavatewan (2023) mukaan datanhallinnan onnistuminen edellyttää, että kaikki datanhallinnan sidosryhmät saadaan osallistumaan datanhallintaan ja vakuuttamaan sen tuottamasta arvosta. Tutkijan mukaan datanhallinnan pitäisi olla koko organisaation yhteinen asia. Tämän toteutumisessa suuressa osassa on se, miten eri sidosryhmät ymmärtävät, arvottavat ja toteuttavat datanhallintaa. Mahanti (2018) jakaa tämän näkemyksen. Hänen mukaansa datanhallinnan sidosryhmien sitouttamiseksi ja muutosvastarinnan vähentämiseksi on tehtävä toimia. Datanhallinta vaikuttaa kaikkiin työntekijöihin, jotka luovat, hallinnoivat kuluttavat, varastoivat ja analysoivat dataa. Tutkijan mukaan kaikki nämä sidosryhmät pitäisi vakuuttaa datanhallinnan tuomasta arvosta. Tämä edellyttää sitä, että organisaatiossa tiedetään, mitä hyötyjä datanhallinta luo. Mahantin (2018) mielestä datanhallinnan edistymistä tulisi seurata muun muassa aikaansaatuja säästöjä tai luodun liikevaihdon kannalta.

Gupta ja Cannon (2020) esittävät tutkimuksessaan, että datanhallinnan määritelmän muodostaminen luo perustan organisaation datanhallinnalle. Heidän mukaansa datanhallintaa ei voi määritellä tyhjiössä, koska datanhallinta on sidottu muun muassa organisaation ominaisuuksiin, kuten strategiaan, prioriteetteihin, visioon, kulttuuriin, organisaatiorakenteeseen ja toimintaympäristöön. Tutkijat korostavat, että datanhallinnan määrittely huolellisesti on olennaista, koska se vaikuttaa siihen, miten työntekijät saadaan hyväksymään ja sitoutumaan datanhallintaan. Työntekijöiden sitoutuminen on tärkeää, koska datanhallinta on koko organisaation vastuulla, ei vain IT-osaston (DAMA 2017; Al-Ruithe et al. 2019; Gupta & Cannon 2020). Jos datanhallinta määritellään epämääräisesti tai se saadaan kuulostamaan vain IT:n asialta, muut sidosryhmät eivät välttämättä pysty näkemään datanhallinnan yhteyttä tai arvoa omalle roolilleen ja se saatetaan tulkita jopa turhaksi byrokratiaksi tai omaa työtä hankaloittavaksi asiaksi. Määritelmän tulisi

siis pystyä tiivistämään mitä organisaatio haluaa datanhallinnalla saavuttaa sekä olla houkutteleva ja merkityksellinen erilaisille sidosryhmille. (Gubta & Cannon 2020) Datanhallinnan määrittelyn ja hyötyjen näkyväksi tuomisen lisäksi työntekijöiden sitoutumista voidaan edistää myös johdon tuen, koulutuksen ja viestinnän kautta. Nämä liittyvät läheisesti muihin tunnistettuihin kriittisiin menestystekijöihin.

3.4.3 Henkilöstön kyvykkyudet

Työntekijöiden datakyvykkyudet ovat keskeinen tehokkaan datanhallinnan mahdollistaja. Datakyvykkyudet kattavat dataan liittyvien toimintojen toteuttamisen ja datanhallinnan roolien ja vastuiden hoitamisen vaatimat tiedot ja taidot sekä ymmärryksen datan ja sen hallinnan tärkeydestä. Datakyvykkyksiä tulisi tarkastella kaikilla organisaatiotasolla, sillä ne kaikki ovat mukana useissa dataan liittyvissä toiminnoissa datan elinkaaren aikana. (Alhassan et al. 2019) Datakyvykkyuksien kehittämiseen ja ylläpitoon liittyy vahvasti viestintä, sillä viestinnällä voidaan lisätä tietoisuutta datanhallinnasta ja sen hyödyistä.

Henkilöstön osaamisen ja kyvykkyuksien merkitys korostuu nopeasti muuttuvissa ympäristöissä. Lillie ja Eybers (2019) väittävät, että datanhallinnassa ja datan operatiivisessa hallinnassa on sisäänrakennettuna tarve ketteryydelle, joka syntyy tarpeesta toimittaa liiketoiminnalle hyödyllistä dataa oikea-aikaisesti, jotta liiketoiminta voi tukea ja mahdollistaa yrityksen liiketoimintastrategian toteutumisen. Tämä ketteryys syntyy kyvystä havaita, sopeutua ja reagoida nopeasti ympäristön muutoksiin. Tutkijoiden mukaan tällaisen ketteryyden saavuttaminen ja ylläpitäminen asettaa vaatimuksia henkilöstön osaamiselle.

3.4.4 Viestintä ja yhteistyö

Datanhallinnan sidosryhmillä tulee olla selkeä ymmärrys siitä, mitä datanhallinta on, miksi se on tärkeää organisaatiolle, mitä muutoksia ollaan tekemässä ja mitä muutokset vaativat sidosryhmiltä. Ihmiset eivät voi muuttaa toimintatapojaan, jos he eivät tiedä, mitä heidän tulisi tehdä eri tavalla. Tehokas kommunikaatio ja viestintä on tärkeässä osassa, jotta työntekijät tietävät mitä heiltä odotetaan ja he voivat työskennellä yrityksen odotusten mukaisesti. (DAMA 2017; Mahanti 2018) Tämän tietoisuuden parantaminen on tärkeä edellytys osallistumisen lisäämiselle, vastarinnan vähentämiselle ja koko organisaation sitouttamiselle datanhallintaan (Abraham et al. 2019). Kun odotukset ja toimintatavat kommunikoidaan tehokkaasti, voidaan varmistaa, että koko yritys toimii samojen periaatteiden mukaisesti. Tässä apuna yritykset voivat käyttää viestintäsuunnitelmaa. Viestintäsuunnitelman tulisi kattaa keinot viestiä miksi ja miten datanhallintaa tehdään ja mitä hyötyjä datanhallinta tuottaa (Mahanti 2018; Abraham et al. 2019).

Sen lisäksi, että datanhallinnan kehittämisestä vastaavan johdon tulisi viestiä datanhallinnan tavoitteista, tärkeydestä ja hyödyistä muulle organisaatiolle, tarvitaan viestintää ja kommunikaatiota myös eri yksiköiden ja roolien välillä. Esimerkiksi liiketoiminnan pitää pystyä viestimään omista datatarpeistaan data- ja IT-ammattilaisille sekä datan omistajien pitää pystyä viestimään datan vaatimuksista datatyön toteuttajille (Alhassan et al. 2019). Bento ja muut (2022) ovat tunnistaneet tehokkaan ja strategian kanssa yhdenmukaisen viestinnän datanhallinnan onnistumisen kannalta kriittiseksi tekijäksi. Tarkoituksenmukainen viestintä mahdollistaa yhteisen päämäärän jakamisen, datanhallinnan tarpeellisuuden ymmärtämisen ja yhteisen kielen luomisen ja näin ollen myös yhteistyön edistämisen (Tsavatewa 2023).

Tehokas viestintä ja kommunikaatio on edellytys tiimien ja yksiköiden väliselle yhteistyölle, joka on myös yksi datanhallinnan kriittisistä elementeistä. Kuten jo aiemmin todettu, datanhallinta vaatii koko organisaation osallistumista ja sitoutumisen aina johdosta liiketoiminnan ja IT:n ammattilaisiin. Erityisesti hajautetussa datanhallinnassa useiden yksiköiden osallistuminen datanhallintaan vaatii tehokasta yhteistyötä. Tiimien välinen yhteistyö edistää omistajuutta, kommunikaatiota ja koko organisaation kattavan kuvan muodostumista. (Mahanti 2018)

3.5 Datanhallinnan arviointi

DAMA:n (2017) mukaan datanhallinnan onnistumista pitäisi seurata ja mitata, jotta yritys voi viestiä ja näyttää datanhallinnan sidosryhmille, kuinka hyvin datanhallinta on luonut liiketoiminta-arvoa ja saavuttanut tavoitteita. Lisäksi olisi tärkeää seurata, onko yrityksellä tarpeeksi resursseja tukea datanhallintaa. DAMA ehdottaa muutamia esimerkkimittareita, jotka liittyvät datanhallinnan arvon, tehokkuuden ja kestävyysmittaamiseen. Heidän mukaansa datanhallinnan arvoa voidaan arvioida muun muassa mittaamalla riskien vähentymistä ja operaatioiden tehokkuuden lisäystä. Datanhallinnan tehokkuutta taas voidaan arvioida mittaamalla tavoitteiden saavuttamista sekä kommunikaation ja koulutuksen tehokkuutta. Datanhallinnan kestävyttä sen sijaan voidaan arvioida muun muassa mittaamalla työntekijöiden standardien ja toimintatapojen noudattamista. DAMA ei kuitenkaan anna tarkempia tietoja siitä, miten näitä käytännössä voidaan mitata. Kuten aiemmin todettu, yrityksillä on erilaiset tavoitteet datanhallinnalle, joten jos halutaan mitata esimerkiksi tavoitteiden täyttymistä, se väistämättä johtaa siihen, että kaikki yritykset eivät voi seurata ja mitata datanhallinnan vaikutuksia samoilla tavoilla.

Aiemmin todettiin, että datanhallinnan täytyy pystyä mukautumaan sisäisiin ja ulkoisiin muutoksiin. Datanhallinnan ja sen vaikutusten seuraaminen ja arviointi voi antaa läpinäkyvyyttä siihen, onko yrityksen datanhallinta vielä ajan tasalla. Jos seurannassa

huomataan, että esimerkiksi datan laatu on selittämättömästi syystä heikentynyt, on se signaali yritykselle, että jotain täytyisi tehdä. Datan käsittelypolitiikkoja tai työntekijöiden osaamista voi esimerkiksi olla tarpeen tarkastella uudestaan.

Koska datanhallinnan suorien vaikutusten mittaaminen voi olla vaikeaa, datanhallinnan tilaa ja edistyneisyyttä voidaan arvioida myös datanhallinnan kypsyysmallien avulla. Kypsyysmallit eivät suoraan tuota tietoa siitä, mitä vaikutuksia datanhallinnalla on, vaan niiden avulla voidaan arvioida yrityksen datanhallinnan tilaa ja kyvykkyyksiä sekä määrittellä, mitä muutoksia täytyy tehdä, jotta kypsyys voidaan parantaa (Herselman et al. 2019). Tekemällä arvioinnin datanhallinnan kypsyudesta lähtötilanteessa ja toisaalta myöhemmin muutosten jälkeen voidaan myös tuottaa tietoa siitä, miten datanhallintaan tehnyt muutokset ovat vaikuttaneet datanhallinnan tilaan.

3.6 Kirjallisuuskatsauksen synteesi

Tämän tutkimuksen kirjallisuuskatsauksessa pyrittiin luomaan käsitys siitä, miten yritys voi muodostaa tai uudistaa datanhallintaa omaan toimintaansa sopivaksi. Kirjallisuuskatsauksen perustana oli aikaisempien tutkimusten tulos siitä, ettei ole olemassa yhtä kaikille yrityksille sopivaa datanhallintamallia, vaan datanhallinta tulisi muodostaa kunkin yrityksen kontekstiin sopivaksi (Weber et al. 2009; Al-Ruithe & Benkhelifa 2017; Sothilingam et al. 2021). Lisäksi katsauksen perustana oli käsitys, jonka mukaan datanhallintaa ei tulisi pitää kertaluontoisena projektina vaan jatkuvana prosessina (DAMA 2017; Walsh et al. 2022), jota pitää pystyä uudistamaan ulkoisten ja sisäisten muutosten mukana (DAMA 2017; Abraham et al. 2019; Mahanti 2019). Näiden näkemysten pohjalta datanhallinnan kehittämisen kannalta merkittäviksi vaiheiksi tunnistettiin organisaatiokohtaisen datanhallintamallin muodostaminen tai uudistaminen, datanhallinnan toteuttaminen, datanhallinnan arviointi sekä toiminnan vakiinnuttaminen tai muutosten tekeminen. Katsauksessa keskityttiin kolmeen ensimmäiseen vaiheeseen.

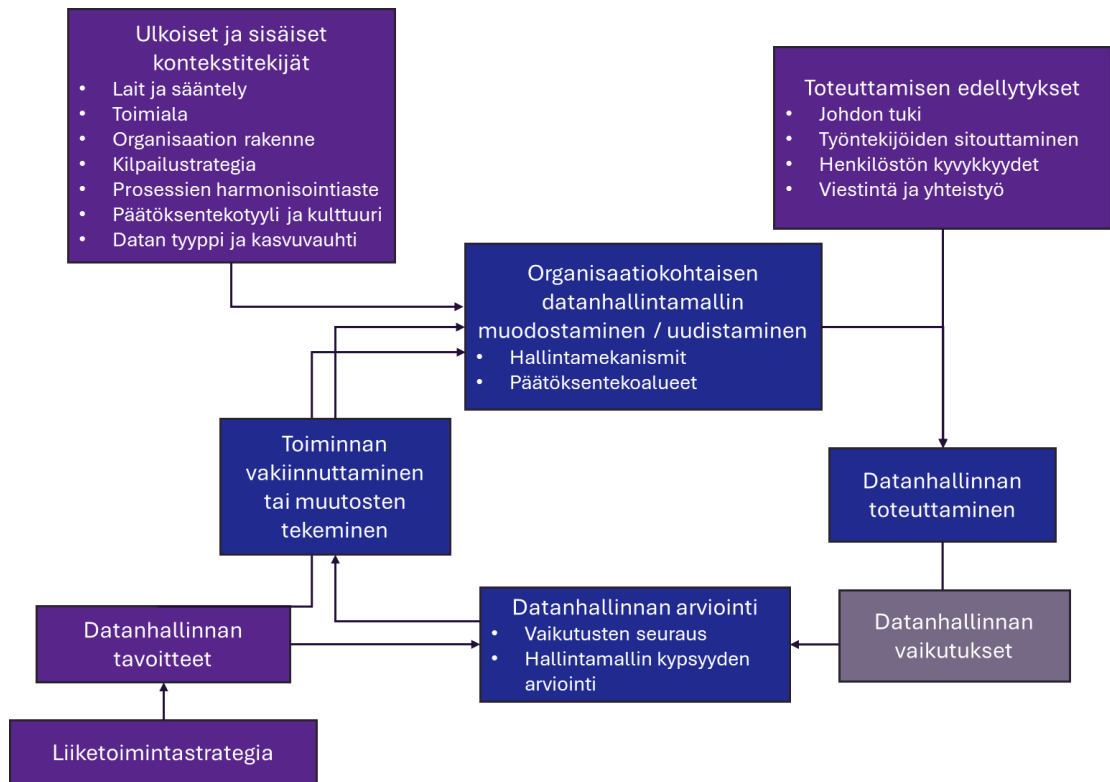
Kirjallisuuskatsaus aloitettiin selvittämällä, mitä datanhallinta on ja mistä osista datanhallintamalli muodostuu. Abrahamin ja muiden (2019) tutkimuksen perusteella datanhallinnasta tunnistettiin kolme tälle tutkimukselle merkittävää osa-aluetta: menettelylliset hallintamekanismit, rakenteelliset hallintamekanismit sekä päätöksentekoaalueet. Menettelyllisissä hallintamekanismeissa tunnistettiin erityisesti organisaation datastrategian ja liiketoimintastrategian yhdistämisen merkitys (Dallemlule & Davenport 2017; Medeiros et al. 2020). Rakenteellisissa hallintamekanismeissa korostui kaikkien strategisten tasojen sisällyttäminen datanhallintaan (Korhonen et al. 2013; DAMA 2017; Chakravorty 2020) sekä datanhallinnan hajautuksen mukauttaminen organisaation tavoitteisiin sopivaksi. Lisäksi organisaatiokohtaisen datanhallintamallin muodostamisessa kriittiseksi tekijäksi

tunnistettiin datanhallintamallin yhteensopivuus yrityksen liiketoimintastrategian ja tavoitteiden (DAMA 2017; Abraham et al. 2019; Chakravorty 2020; Lis et al. 2023) sekä muiden sisäisten ja ulkoisten kontekstitekijöiden kanssa (Weber et al. 2009; Otto 2011b; Abraham et al. 2019; Lis et al. 2023).

Datanhallinnan toteuttamisessa kriittisiksi tekijöiksi tunnistettiin johdon tuki (DAMA 2017; Mahanti 2018; Al-Ruithe et al. 2019), työntekijöiden sitouttaminen datanhallintaan (Al-Ruithe et al. 2019; Karkošková 2023), henkilöstön kyvykkyydet (Mahanti 2018; Al-Ruithe et al. 2019) sekä viestintä ja yhteistyö (DAMA 2017; Mahanti 2018; Bento et al. 2022). Johdon tuki tunnistettiin oleelliseksi tekijäksi datanhallinnan onnistuneen toteuttamisen kannalta, sillä datanhallinnan kehittäminen vaatii usein organisaatiotason muutoksia, joiden tekemiseen tarvitaan johdon hyväksyntä (Mahanti 2018; Chakravorty 2020). Henkilöstön sitouttaminen datanhallintaan nousi kriittiseksi tekijäksi, sillä datanhallinta vaikuttaa moniin työntekijöihin ja heidän toimintaansa. Jotta datanhallinta voi olla koko organisaation eikä vain IT-osaston asia, henkilöstön täytyy ymmärtää sen arvo ja oma roolinsa sen edistämiseksi. (Mahanti 2018; Tsavatewan 2023) Työntekijöiden sitoutumisen lisäksi datanhallinnan toteuttamisen kannalta oleelliseksi tekijäksi tunnistettiin henkilöstön kyvykkyydet. Keskeisiksi kyvykkyyksiksi tunnistettiin dataan liittyvien toimintojen toteuttamisen ja datanhallinnan roolien ja vastuiden hoitamisen vaatimat tiedot ja taidot sekä ymmärrys datan ja sen hallinnan tärkeydestä (Alhassan et al. 2019). Viimeisiksi datanhallinnan toteuttamisen kannalta kriittisiksi tekijöiksi tunnistettiin viestintä ja yhteistyö. Tehokasta kommunikaatiota ja viestintää tarvitaan, jotta työntekijät tietävät mitä heiltä odotetaan ja he voivat työskennellä yrityksen odotusten mukaisesti (DAMA 2017; Mahanti 2018). Tiimien välistä yhteistyötä taas tarvitaan, sillä se edistää omistajuutta, kommunikaatiota ja koko organisaation kattavan kuvan muodostumista (Mahanti 2018).

Kirjallisuuskatsauksen lopuksi todettiin vielä datanhallinnan arvioinnin oleellisuus. Datanhallinnan onnistumisen seuranta todettiin tärkeäksi, jotta sidosryhmille voidaan viestiä ja näyttää, kuinka hyvin datanhallinta on luonut liiketoiminta-arvoa ja saavuttanut tavoitteita (DAMA 2017). Tällä voidaan lisätä työntekijöiden sitoutuneisuutta datanhallintaan. Sopivien mittareiden todettiin kuitenkin riippuvan pitkälti yrityksen datanhallinnan tavoitteista, jotka voivat vaihdella suuresti organisaatioiden kesken. Viimeisenä vielä todettiin, että datanhallintaa kehittäessä datanhallinnan tilaa ja kypsyyttä voidaan arvioida datanhallinnan kypsyyksille. Kypsyyksimallit auttavat suunnittelemaan datanhallinnan kehitystoimenpiteitä ja arvioimaan niiden vaikutusta (Herselman et al. 2019).

Kirjallisuuskatsauksen lopputuloksena muodostettiin malli, joka kuvaa, miten datanhallintaa voidaan muodostaa tai uudistaa organisaation toimintaan sopivaksi. Malli on kuvattu kuvassa 10.



Kuva 10. Kirjallisuuskatsauksen perusteella luotu malli datanhallinnan jatkuvaan kehittämiseen

Kehitys on kuvattu ympyrässä, koska kirjallisuuskatsauksessa todettiin, että datanhallinta on jatkuva prosessi, jonka tulee mukautua sisäisiin ja ulkoiisiin muutoksiin (DAMA 2017; Walsh et al. 2022). Se vaatii siis jatkuvaa kehittämistä ja parantamista. Kehitysprosessin vaiheiden lisäksi mallissa on kuvattu tekijöitä, jotka vaikuttavat prosessin vaiheisiin ja jotka tulisi huomioida, jotta datanhallinta tukisi organisaation toimintaa mahdollisimman hyvin.

Datanhallinnan kehittäminen alkaa organisaatiokohtaisen datanhallintamallin muodostamisesta tai uudistamisesta. Sen perustana on organisaation liiketoimintastrategian ja tavoitteiden johtaminen datanhallinnan tavoitteiksi sekä ulkoisten ja sisäisten kontekstitekijöiden arviointi. Tämän jälkeen datanhallintamalli muodostetaan tai mukautetaan näihin tekijöihin sopivaksi. Organisaatiokohtaisen datanhallintamallin muodostamisen tai uudistamisen jälkeen malli otetaan käyttöön ja datanhallintaa toteutetaan. Organisaation tulisi kuitenkin samalla arvioida, onko datanhallinnan toteuttamisen edellytykset kunnossa eli onko datanhallinnan toteuttamiseen vaikuttavat tekijät, kuten työntekijöiden kyvykkyydet, organisaatiossa tarvittavalla tasolla vai tarvitaanko toimenpiteitä. Kun datanhallintaa toteutetaan, syntyy vaikutuksia, joita tulisi arvioida. Ainakin osa mittareista olisi hyvä johtaa datanhallinnan tavoitteista, jotta voidaan arvioida, kuinka hyvin datanhallinnalla on saavutettu sille asetettuja tavoitteita. Arvioinnin perusteella datanhallinta voidaan vakiinnut-

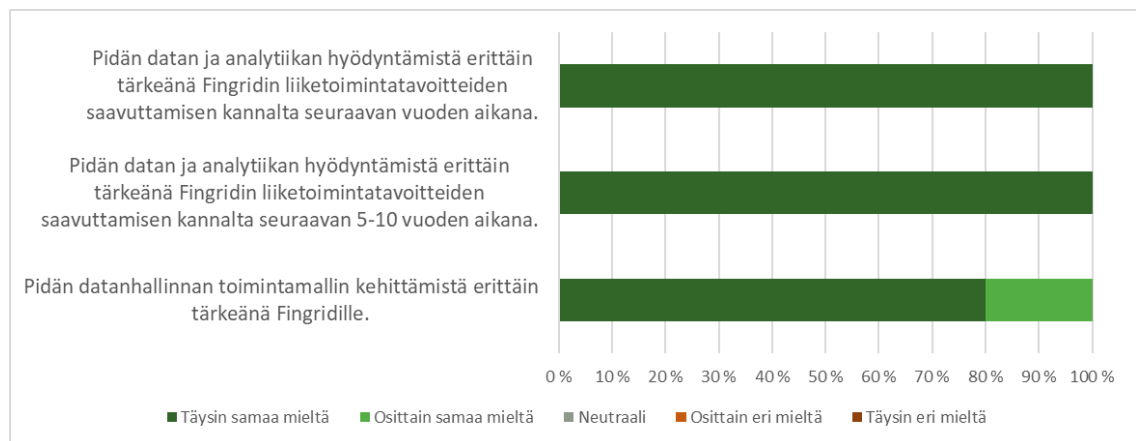
taa käyttöön tai todeta tarve muutoksille. Kuten edellä todettu, datanhallinnan kehittäminen ei kuitenkaan ole kertaluontoinen tapahtuma, sillä datanhallinta on jatkuva prosessi, jonka tulisi olla yhteensopiva organisaation muuttuvien tavoitteiden ja ulkoisten ja sisäisten kontekstitekijöiden kanssa. Näin ollen sen kehittämisen ja arvioinnin tulisi olla jatkuvaa.

4. EMPIIRISEN TUTKIMUKSEN TULOKSET

Tässä luvussa kuvataan haastateltaville lähetetyn taustoittavan kyselyn sekä haastatteluiden tulokset.

4.1 Taustoittavan kyselyn tulokset

Kyselyssä kartoitettiin vastaajien näkemyksiä datan, analytiikan ja datanhallinnan merkityksestä. Kaikki vastaajat organisaatiotasosta riippumatta pitivät datan ja analytiikan hyödyntämistä erittäin tärkeänä kohdeorganisaation liiketoimintatavoitteiden saavuttamisen kannalta niin pitkällä kuin lyhyellä aikavälillä. Datan ja analytiikan merkityksen lisäksi vastaajat tunnistavat sen hallinnan oleellisuuden. Kaikki vastaajat olivat joko osittain tai täysin samaa mieltä datanhallinnan toimintamallin kehittämisen tärkeydestä (kuva 11).



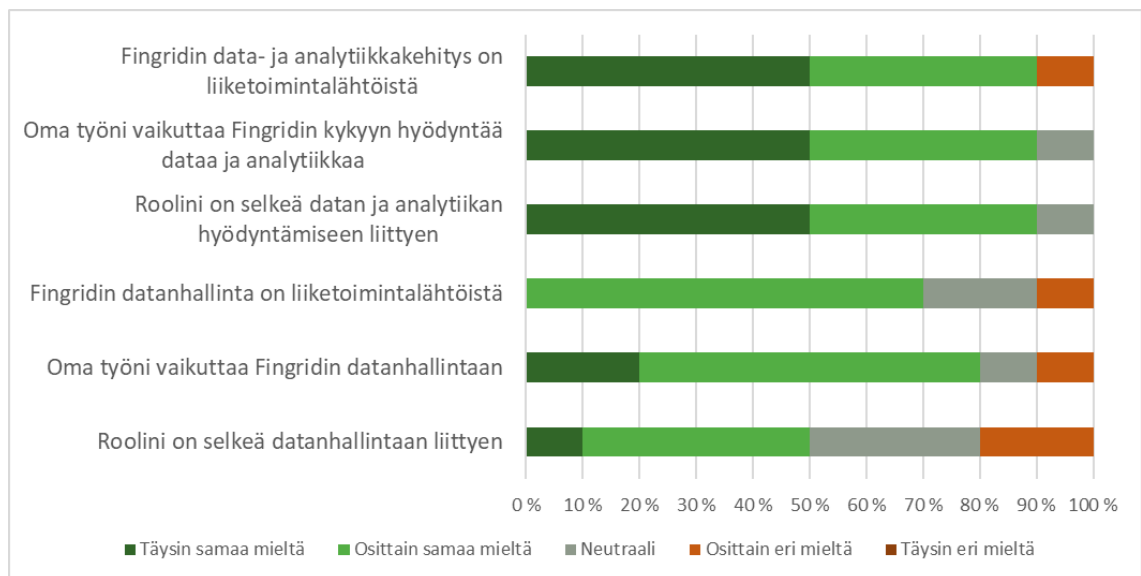
Kuva 11. Datan ja analytiikan hyödyntämisen sekä datanhallinnan tärkeys

Seuraavaksi kartoitettiin, mihin asioihin vastaajat liittävät datanhallinnan mahdollisuudet. Tärkeimmiksi datanhallinnan mahdollistamiksi asioiksi nousi toiminnan tai prosessien optimointi ja ennustekyvyn kehittäminen sekä strategian toimeenpano. Viimeisimmäksi jäi riskien vähentäminen (taulukko 8).

Taulukko 8. *Datanhallinnan mahdollisuudet*

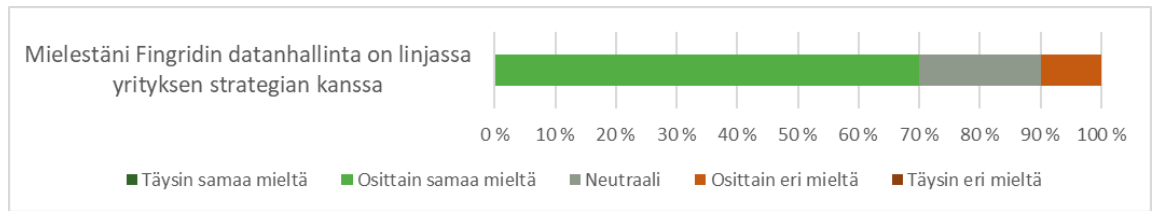
	Ensimmäinen valinta	Toinen valinta	Kolmas valinta	Neljäs valinta
Toiminnan / prosessien optimointi ja ennustekyvyn kehittäminen	4	3	1	0
Strategian toimeenpano	3	2	1	2
Toiminnan / prosessien seuraaminen ja raportointi	1	2	4	1
Riskien vähentäminen, esim. toiminnan säännöstenmukaisuuden varmistamisen kautta	0	1	2	5

Vastaajilta haluttiin näkemyksiä myös data- ja analytiikkakehityksen ja datanhallinnan liiketoimintalähtöisyydestä. Data- ja analytiikkakehityksen vastaajat kokivat pääosin liiketoimintalähtöiseksi. Vastaajista 50 % koki sen täysin liiketoimintalähtöisenä ja vain yksi vastaaja oli asiasta osittain eri mieltä. Vastaajat myös tunnistivat, että heidän roolillaan on vaikutusta kohdeorganisaation kykyyn hyödyntää dataa ja analytiikkaa ja tämä rooli oli melko selkeä. Datanhallinnan liiketoimintalähtöisyys ja vastaajien rooli siihen liittyen ei ollut aivan yhtä selkeä. Kukaan vastaajista ei ollut täysin samaa mieltä datanhallinnan liiketoimintalähtöisyydestä, mutta kuitenkin 70 % oli asiasta osittain samaa mieltä. Vastaajista 80 % koki roolillaan olevan vaikutusta datanhallintaan, mutta vain yhden vastaajan mielestä tämä rooli oli täysin selkeä (kuva 12). Vastaajat tunnistavat siis roolinsa vaikuttavan datanhallintaan, mutta vastausten perusteella ei ole täysin selkeää, mitä kaikkea tämä rooli sisältää.



Kuva 12. Liiketoimintalähtöisyys ja roolien selkeys

Vastaajia pyydettiin myös arvioimaan datanhallinnan yhteyttä kohdeorganisaation strategiaan. Vastaajista 70 % oli osittain samaa mieltä siitä, että yrityksen datanhallinta on linjassa organisaation strategian kanssa (kuva 13). Täysin selkeä tämä yhteys ei siis kuitenkaan ollut.



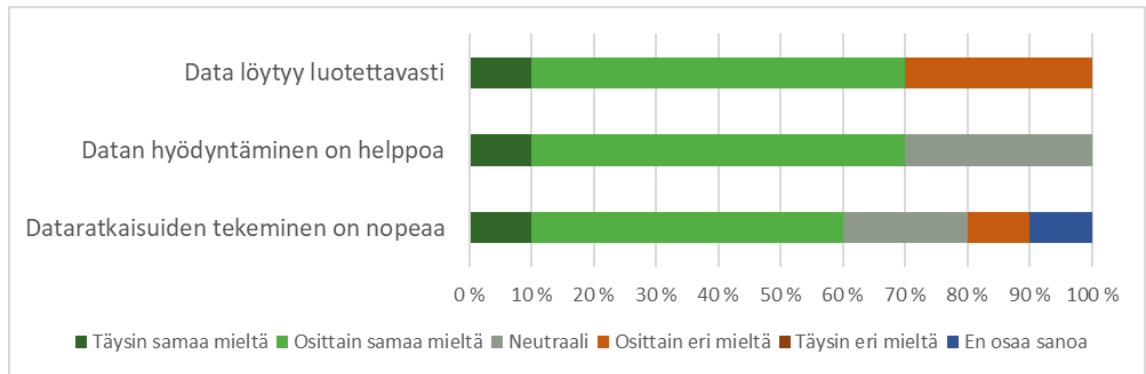
Kuva 13. Datanhallinnan yhteys kohdeorganisaation strategiaan

Tämän jälkeen vastaajia pyydettiin arvioimaan toimintamalleissa usein vastakkaisina pareina nähtävien tekijöiden tärkeyttä datan hyödyntämiseen ja hallintaan liittyen. Taulukosta 9 nähdään, että tekijöistä erottui selkeästi datan tehokkaan hyödyntämisen ja joustavuuden merkitys. Vastaajista 80 % valitsi datan tehokkaan hyödyntämisen datan turvaamisen sijaan ja 90 % valitsi joustavuuden kontrollin sijaan. Säännöstenmukaisuuden ja innovatiivisuuden välillä ei syntynyt yhtä selkeää jakoa. Puolet vastaajista valitsi säännöstenmukaisuuden ja puolet vastaajista innovatiivisuuden.

Taulukko 9. *Datan hyödyntämiseen ja hallintaan liittyvien tekijöiden tärkeys kohdeorganisaation liiketoiminnalle*

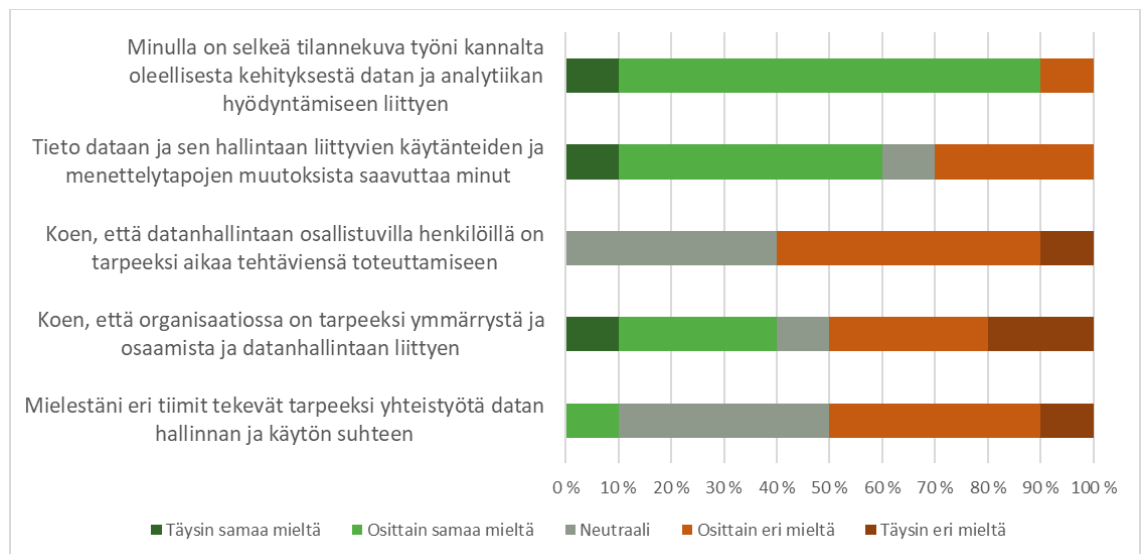
	Vaihtoehto A selkeästi tärkeämpi	Vaihtoehto A hieman tärkeämpi	Vaihtoehto B hieman tärkeämpi	Vaihtoehto B selkeästi tärkeämpi	En osaa sanoa
A) datan turvaaminen vai B) datan tehokas hyödyntäminen	0	1	3	5	1
A) kontrolli vai B) joustavuus	0	1	7	2	0
A) säännöstenmukaisuus vai B) innovatiivisuus	0	5	4	1	0

Lisäksi vastaajia pyydettiin arvioimaan kohdeorganisaation datanhallinnan vision toteutumista (kuva 14). Vision kolmesta osasta haasteita voidaan tunnistaa erityisesti datan löydettävyydessä. Vastaajista 30 % oli osittain eri mieltä siitä, että data löytyy luotettavasti. Datan hyödyntämisen helppoudesta ei varsinaisesti oltu eri mieltä, mutta vain yksi vastaaja oli väitteen kanssa täysin samaa mieltä. Dataratkaisuiden tekemistä ei nähty erityisen nopeana, sillä vain yksi vastaaja vastasi olevansa täysin samaa mieltä siitä, että dataratkaisuiden tekeminen on nopeaa.



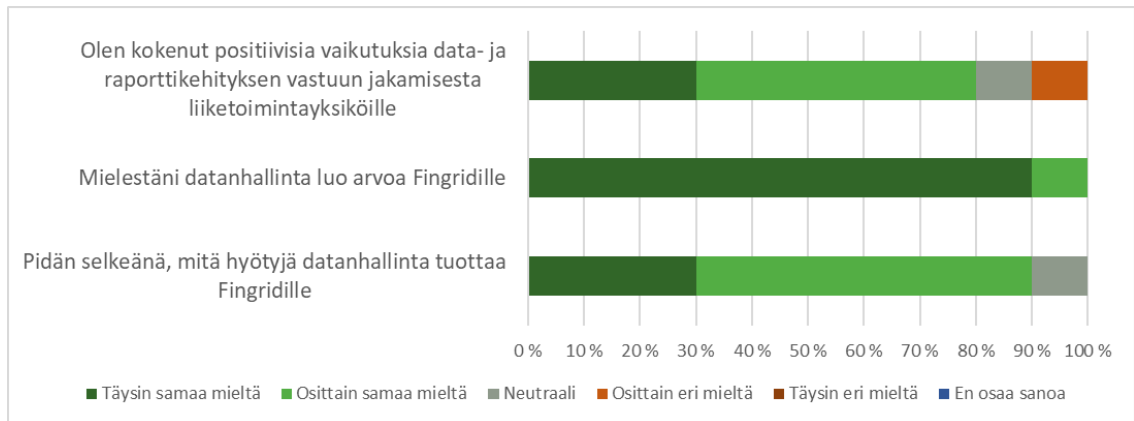
Kuva 14. Datanhallinnan vision toteutumisen

Seuraavaksi kyselyssä kartoitettiin datanhallinnan toteuttamiseen liittyvien tekijöiden tilaa (kuva 15). Vastauksista nousee selkeästi esille datanhallinnan resurssien riittämättömyys, datanhallintaan liittyvän osaamisen taso sekä tiimien välinen yhteistyö. Yksikään vastaajista ei ollut samaa mieltä siitä, että datanhallintaan osallistuvilla henkilöillä olisi riittävästi aikaa tehtäviensä toteuttamiseen. Vastaajista 50 % oli osittain tai täysin eri mieltä siitä, että organisaatiossa on tarpeeksi ymmärrystä ja osaamista datanhallintaan liittyen. Sama tilanne on myös tiimien välisen yhteistyön kanssa.



Kuva 15. Datanhallinnan toteuttamiseen liittyvien tekijöiden tila

Viimeiseksi kyselyssä kartoitettiin vastaajien näkemyksiä hajautetun data- ja raporttikehityksen sekä datanhallinnan vaikutuksista (kuva 16).



Kuva 16. Hajautetun data- ja raporttikehityksen sekä datanhallinnan vaikutukset

Vastauksista nähdään, että hajautettu data- ja raporttikehitys koetaan selkeästi positiivisena asiana. Vastaajista 80 % oli kokenut positiivisia vaikutuksia data- ja raporttikehityksen vastuun jakamisesta liiketoimintayksiköille. Datanhallinnan arvosta kohdeorganisaatiolle vastaajat olivat yhtä mieltä. Kaikkien vastaajien mielestä datanhallinta luo arvoa kohdeorganisaatiolle. Datanhallinnan tuottamat hyödyt eivät kuitenkaan olleet täysin selkeitä. Vastaajista vain 30 % oli täysin samaa mieltä datanhallinnan tuottamien hyötyjen selkeydestä.

Kyselyllä saatiin kartoitettua yrityksen datanhallinnan nykytilaa selkeästi koostettavissa olevassa muodossa. Taustoittavan kyselyn tulosten visualisoinneista nähdään nopeasti, missä datanhallinnan osa-alueissa haastateltavat näkevät haasteita ja mihin kohdeorganisaation tulisi kiinnittää huomiota. Kyselyn avulla oli myös helppoa verrata eri organisaatiotasojen vastauksia. Organisaatiotasojen välillä ei kuitenkaan tunnistettu suuria näkemuseroja. Saman organisaatiotason edustajat saattoivat olla asioista eri mieltä tai samaa mieltä. Joitakin huomioita yhdistävistä näkemyksistä nostettiin kuitenkin seuraavassa kappaleessa haastatteluiden tulosten kuvaamisen yhteydessä.

Kohdeorganisaation datanhallinnan nykytilan kartoittaminen oli tutkimuksen kannalta oleellista, sillä sen avulla voidaan tunnistaa, mitä osa-alueita kohdeorganisaation olisi tärkeää kehittää ja mitä vahvuuksia sen tulisi ylläpitää. Näitä ja haastatteluiden löydöksiä voidaan peilata teoreettisen viitekehityksen malliin ja arvioida, voitaisiinko mallia hyödyntää kohdeorganisaation datanhallinnan uudistamisessa ja saada näin yrityksen datanhallinta tukemaan vielä vahvemmin sen toimintaa.

4.2 Haastatteluiden tulokset

Haastatteluaineiston teemoittelun tuloksena aineistosta tunnistettiin 16 teemaa ja 5 näitä yhdistävää pääteemaa. Teemat on kuvattu taulukossa 10. Seuraavissa alaluvuissa on esitelty teemoja tarkemmin sekä avattu niiden avulla haastatteluiden sisältöä.

Taulukko 10. Haastatteluaineistosta tunnistetut teemat

Pääteema	Teemat
Liiketoimintalähtöisyys	<ul style="list-style-type: none"> Datan arvo liiketoiminnalle Data- ja analytiikkakehityksen liiketoimintalähtöisyys Datanhallinnan liiketoimintalähtöisyys
Toimintamalleihin vaikuttavat tekijät	<ul style="list-style-type: none"> Toimintaympäristön nopea muuttuminen Toimialan säännökset ja vaatimukset Yrityksen toiminnan laajuus Päätöksentekokulttuuri
Datanhallinnan rakenteet	<ul style="list-style-type: none"> Datanhallinnan roolit ja vastuut Määritellyt toimintatavat
Datanhallinnan toteutus	<ul style="list-style-type: none"> Datanhallinnan jalkauttaminen Yhteistyö Resurssit Osaaminen Teknologiat
Datanhallinnan hyödyt	<ul style="list-style-type: none"> Tunnistetut datanhallinnan hyödyt Hyötyjen konkretisointi

4.2.1 Liiketoimintalähtöisyys

Liiketoimintalähtöisyys muodostettiin pääteemaksi yhdistämään teemat ”Datan arvo liiketoiminnalle”, ”Data- ja analytiikkakehityksen liiketoimintalähtöisyys” sekä ”Datanhallinnan liiketoimintalähtöisyys”. Taulukossa 11 on esitelty teemojen muodostuminen.

Taulukko 11. Teeman ”Liiketoiminnan tarpeet” muodostuminen

Pääteema	Teemat	Koodit	Mainittu haastatteluissa
Liiketoimintalähtöisyys	Datan arvo liiketoiminnalle	Datan hyödyntäminen toiminnan pohjana	H5, H6, H7, H8, H9, H10
		Data automatiikassa	H5
		Data tilannekuvan luojana	H6, H7, H10
	Data- ja analytiikkakehityksen liiketoimintalähtöisyys	Tarpeeseen tehdyt dataratkaisut	H1, H2, H4, H6, H7, H8, H9, H10
		Nopeat dataratkaisut	H3, H7, H9
	Datanhallinnan liiketoimintalähtöisyys	Liiketoiminnan yhteys datanhallintaan	H2, H5, H10
Strategian yhteys datanhallintaan		H1, H3, H5	

Aineistosta tunnistettiin osia, joissa tuotiin esille datan hyödyntämisen merkitys liiketoiminnalle sekä datan tärkeitä käyttötarkoituksia. Näistä muodostettiin teema ”Datan arvo liiketoiminnalle”. Haastateltavat tunnistivat datan keskeisyyden omissa toiminnoissaan, mutta myös laajemmin koko organisaation tasolla. Monet haastateltavista mainitsivat yrityksen toiminnan perustuvan pitkälti dataan ja sen hyödyntämiseen (H5, H6, H7, H8, H9, H10). Osa vastaajista korosti, että datalla on arvoa vain silloin, kun sitä hyödynnetään

(H7, H9, H10). Datan merkitys tunnistettiin esimerkiksi automaattiprosessien raaka-aineena (H5) sekä tilannekuvan mahdollistajana (H6, H7, H10).

H10: *"Data on yksi osa tätä johtamista, että se on ihan samanlainen assetti kuin meidän fyysiset assetit, että täytyy olla kunnossa."*

H8: *"Ilman dataa meillä ei (omassa toiminnossa) ole yhtään mitään. Ei täällä mitään fyysistä rakenneta tai muuta, me pohjataan siihen mitä mittaustietoa tuolta verkosta tulee ja mitä ennusteita me luodaan ja mitä päätöksiä me osataan niiden ennusteiden perusteella tehdä..."*

H10: *"Jos meillä on hyvää dataa, joka me hillotaan vaan johonkin purkkiin, josta sitä ei voi sitten oikein käyttää, niin sitten sen datan arvo on aika vähäinen, koska kyllähän se datan arvo tulee siitä, että mitä sillä voidaan tehdä ja kuinka sitä voidaan hyödyntää. Jos ei sitä mihinkään voida käyttää, niin mun mielestä sillä datalla ei ole silloin mitään arvoa."*

Haastatteluaineiston maininnat datan hyödyntämisen tärkeydestä sen arvonluonnin kannalta selittävät kyselyn tulosta, jonka mukaan 80 % vastaajista valitsi datan tehokkaan hyödyntämisen sen turvaamista tärkeämmäksi asiaksi. Kohdeorganisaatiossa on selkeästi eri organisaatiotasot kattava näkemys datan arvosta liiketoiminnalle sekä ymmärrys sen tehokkaan hyödyntämisen tarpeellisuudesta.

Datan arvoon liittyvien osioiden lisäksi haastatteluaineistosta tunnistettiin mainintoja liiketoiminnan tarpeista ja niiden täyttymisestä data- ja analytiikkakehitykseen liittyen. Näistä maininnoista luotiin teema "Data- ja analytiikkakehityksen liiketoimintalähtöisyys". Haastatteluissa tuotiin esille, että data- ja analytiikkakehitystä tehdään liiketoimintayksiköissä, minkä takia data- ja raporttikehittäjillä on ymmärrystä liiketoiminnasta ja sen tarpeista. Haastateltavat toivat esille, että he pitävät toimintatapaa paljon tehokkaampana, kuin esimerkiksi IT-yksikölle keskitettyä data- ja analytiikkakehitystä (H2, H8). Nyt kehitystarpeiden nähtiin tulevan suoraan liiketoiminnalta (H1, H2, H3, H4, H6, H7, H8, H9, H10).

H2: *"Analytiikkaa tehdään liiketoimintayksiköissä, eikä meillä ole mitään keskitettyä IT organisaatiossa olevaa dataa ja analytiikkapalvelua. Ne on siellä liiketoiminnassa sisällä ne tyypit, jotka tekee raporttikehitystä. Jos olisi joku keskitetty IT:n data analytiikkapalvelu, niin siinä on aina semmoinen kuilu, että ne ei ymmärrä, että mitä se liiketoiminta oikeasti tarvitsee. Ihan sama mitkä palaverit käydään ja kerrotaan ja näin edelleen niin se ei sitten välity sinne kumminkaan. Se on ihan eri juttu, kun itse työskentelet siellä liiketoiminnassa."*

Lisäksi hajautetun raporttikehityksen etuna nähtiin sen nopeus.

H7: *"Se on paljon nopeampaa, kun pääsee itse päättämään, että minkälainen se raportti on ja saa teetettyä sen sitten asiantuntijoilla semmoiseksi, kun haluaa."*

Eräs haastateltava nosti kuitenkin esiin, että vaikka data- ja analytiikkakehitystä tehdään liiketoiminnan tarpeisiin, se ei ole aina kovin liiketoimintavetoista.

H1: *"Liiketoiminnalle tätä tehdään tosiaan, mutta aika usein se on niin päin, että datakehitys, raporttikehitys ja tuoteomistaja kontaktoi liiketoimintaa eikä toisinpäin... tuntuu, että olen välillä niin kuin puhelinmyyjä, joka soittelee, että teille olisi tämmöistä tulossa, kiinnostaako."*

Haastatteluaineistosta tunnistettiin myös mainintoja datanhallinnan yhteydestä liiketoiminnan tarpeisiin ja prioriteetteihin sekä kohdeorganisaation strategiaan. Näistä muodostettiin teema "Datanhallinnan liiketoimintalähtöisyys". Aineistosta ei tullut ilmi yhtä vahvaa yhteyttä liiketoiminnan ja datanhallinnan välillä kuin liiketoiminnan ja data- ja analytiikkakehityksen välillä. Tätä tulosta tukevat taustoittavan kyselyn vastaukset, joissa datanhallinnan liiketoimintalähtöisyyttä ei vahvasti kyseenalaistettu (70 % osittain samaa mieltä), mutta se ei myöskään ollut samalla tasolla kuin data- ja analytiikkakehityksen liiketoimintalähtöisyys (90 % täysin tai osittain samaa mieltä). Haastatteluiden vastausten perusteella datanhallinnan nähtiin olevan enemmän ICT-vetoista (H2), eikä datanhallinnan myöskään nähty olevan liiketoiminnan erityinen kiinnostuksen kohde (H2, H5).

H2: *"Datanhallinta ei ratkaise itsessään yhtäkään liiketoiminnallista ongelmaa, vaan se tarjoaa ympäristön saada aikaan yhdenmukaisia hallittavia ratkaisuja. Mutta ei se ratkaise yhtäkään konkreettista ongelmaa itsessään ja sen takia se on jo kaukana liiketoiminnan substanssiasiantuntijan näkökulmasta ja se koetaan sellaisena byrokratiana helposti ja sen hyötyjä on vaikea nähdä, jos ei ole kiinnostusta perehtyä asiaan."*

H5: *"Liiketoiminnassa on monesti ehkä vähän tärkeämpää päästä asioissa eteenpäin. Me vähän välillä ehkä unohdetaankin, kuinka tärkeätä se niitten raamittaminen ja tämän tyylinen homma olisi. Nää kaksi siinä vähän aina tasapainottelee, että pannaanko me malleja kuntoon ja mietitään rakenteellisuutta vai pannaanko vaan dataa liikkumaan ja saadaan homma eteenpäin."*

Toisaalta yksi haastateltava näki datanhallinnan vaatimusten syntyvän liiketoiminnan tarpeista, vaikka konkreettinen tekeminen ei olekaan liiketoiminnasta lähtöisin.

H10: *"Liiketoiminnan tarpeethan sitä kautta pelkistyy siihen, että mitä vaatimuksia ne asettaa datanhallinnalle, että onhan siinä mielessä sekin omalla tavallaan liiketoimintalähtöistä. Mutta jos me ajatellaan konkreettista tekemistä, niin se ei tule suoraan tietenkään liiketoiminnasta. Ei, että hei nyt ruvetaan tekemään uutta tietoallasta, vaan kyllähän se tulee sitten sieltä, että todetaan, että kun on näitä (liiketoiminnan) tarpeita niin jotta*

me pystytään toteuttamaan ne, se edellyttää, että meidän data on vähän paremmin organisoitua kuin mitä se on ollut.”

Aineiston maininnat datanhallinnan ja organisaation strategian yhteydestä liittyivät pitkälti strategian toteutumisen edellytyksiin ja niiden mahdollistamiseen. Strategian toteutumisen nähtiin vaativan saatavilla olevaa, hyvälaatuista dataa, jota datanhallinta voi mahdollistaa (H3, H6).

H3: *”Tokihan meillä on strategiassa paljon mihin tarvitaan dataa, niin sitten sen mahdollistaminen, että meillä on myös se saatavilla.”*

Osa haastateltavista ei kuitenkaan nähnyt vahvaa yhteyttä strategian ja datanhallinnan välillä.

H5: *”Ristiriitaisuutta en näe, mutta en myöskään näe, että otetaan täysin kaikki irti mitä voitaisiin.”*

Tämä näkökulma jaetaan kohdeorganisaatiossa todennäköisesti laajemminkin. Taustoitettavan kyselyn tulokset näyttävät, että suurin osa vastaajista on osittain samaa mieltä siitä, että kohdeorganisaation datanhallinta on linjassa sen strategian kanssa. Tämän yhteyden teroittaminen voisi lisätä organisaatiossa ymmärrystä siitä, miksi datanhallintaa tehdään ja miten se tuottaa arvoa laajemmassa mittakaavassa.

Haastatteluiden perusteella liiketoiminnan ja liiketoiminnan tarpeiden merkitys on oleellinen osa kohdeorganisaation data- ja analytiikkakehitystä. Kohdeorganisaatiossa vaikuttaa olevan myös laajemminkin selkeä halu tehdä asioita liiketoiminnan tarpeisiin. Datanhallinnassa tämä liiketoimintalähtöisyys ei kuitenkaan ole yhtä selkeästi havaittavissa kuin data- ja analytiikkakehityksessä. Tämä on siinä mielessä luonnollista, että teoreettisen viitekehyksen (kuva 15) mukaan datanhallintaa ei kehitetä suoraan liiketoiminnan tarpeiden kautta, vaan liiketoiminnan tavoitteet ja tarpeet pelkistyvät ensiksi datanhallinnan tavoitteiksi. Eli kuten H10 totesi, liiketoiminnan tarpeet eivät suoraan ohjaa datanhallinnan suuntaa, mutta ne asettavat vaatimuksia datanhallinnalle. Tämä ei kuitenkaan välity yhtä selkeänä kytköksenä kuin esimerkiksi raporttikehityksessä, jossa liiketoiminta määrittelee suoraan, millaisen raportin he haluavat ja mihin tarpeeseen. Kuitenkin linjaus viedä data- ja analytiikkakehitys lähemmäs liiketoimintayksiköitä on datanhallinnasta peräisin. Eli kohdeorganisaatiossa on selkeästi pyrkimys tehdä myös datanhallintaa liiketoimintalähtöisesti. Tämä on tärkeää, sillä datanhallinnan rakenteet ja linjaukset vaikuttavat liiketoimintaan. Jos datanhallintaa kehitetään välittämättä liiketoiminnan tarpeista, voidaan päätyä luomaan käytänteitä, jotka koetaan turhiksi ja byrokraattisiksi.

4.2.2 Toimintamalleihin vaikuttavat tekijät

Toimintamalleihin vaikuttavat tekijät muodostettiin pääteemaksi yhdistämään teemat ”Toimintaympäristön nopea muuttuminen”, ”Toimialan säännökset ja vaatimukset”, ”Yrityksen toiminnan laajuus” sekä ”Päätöksentekokulttuuri”. Taulukossa 12 on esitetty teemojen muodostuminen.

Taulukko 12. Teeman ”Toimintamalleihin vaikuttavat tekijät” muodostuminen

Pääteema	Teemat	Koodit	Mainittu haastateluisissa
Toimintamalleihin vaikuttavat tekijät	Toimintaympäristön nopea muuttuminen	Toimintaympäristön muutosvauhti	H5, H6, H7, H9
	Toimialan säännökset ja vaatimukset	Yleiset säännökset	H7, H9
		Toimialan vaatimukset	H3, H4, H5, H6
	Yrityksen toiminnan laajuus	Henkilöstön määrä	H6, H10
		Datan määrä	H6, H10
	Päätöksentekokulttuuri	Suhtautuminen päätöksentekorakenteisiin	H1, H2, H3, H4, H8, H10

Aineistosta tunnistettiin mainintoja toimintaympäristöjen nopeasta muuttumisesta ja sen vaikutuksista kohdeorganisaatioon. Toimintaympäristön nopean muuttumisen nähtiin asettavan vaatimuksia yrityksen yksiköiden ja tiimien väliselle yhteistyölle (H6) sekä toimintamallien joustavuudelle (H5, H7, H9). Näistä maininnoista muodostettiin teema ”Toimintaympäristön nopea muuttuminen”.

H6: *”Vanhassa maailmassa riitti, että jokainen hoiti oman tonttinsa. Nyt se ei riitä enää vaan joutuu tarkemmin katsoo. Tää maailma menee vaan niin mahdotonta vauhtia eteenpäin ja tää vihreä murros, niin kyllä se haastaa meitäkin sen yhteistoiminnan kanssa.”*

H7: *”Toimintaympäristöt, toimintamallit ja käytännöt muuttuu nykyään tosi nopeasti, niin mun mielestä olisi tosi tärkeitä, että meidän toimintamallit ja toimintatavat ja myös raportointi pystyy sitten muuttumaan nopeasti tässä muuttuvassa maailmassa.”*

Tätä tarvetta muuntautumiskyvylle ja joustavuudelle tukee myös taustoittavan kyselyn tulokset, joissa vastaajista 90 % valitsi joustavuuden kontrollia vähintään hieman tärkeämmäksi asiaksi. Haastattelussa kysyttäessä kaikki haastateltavat kokivat, että yleisesti ottaen yhtiön toimintamallit ovat tarpeeksi joustavia.

Teema ”Toimialan säännökset ja vaatimukset” syntyi yhdistämällä mainintoja yleisistä säännöksistä ja kohdeorganisaation toimialalla vaikuttavista vaatimuksista. Mainintojen perusteella haastateltavat näkivät toimialan vaatimusten vaikuttavan yrityksen toimintaan ja toimintamalleihin. Haastateltavat nostivat esiin, että yrityksen toimiala on melko säänneltyä (H3, H4, H6) ja toiminnan luonteen vuoksi riskien pienentäminen on tärkeässä osassa (H5), mikä rajoittaa yrityksen toimintaa ja esimerkiksi datan innovatiivista

hyödyntämistä. Säännösten ja ohjeiden noudattamista korostettiin kuitenkin tärkeänä asiana (H7, H9).

H5: *”Me ollaan täällä riskit, riskit ja sähköt päällä, sähköt päällä ympäristössä. Niin meidän innovatiivisuus on aina pikkaisen rajoitetumpaa monesti, kun ehkä jossakin muun tyyllisessä bisneksessä. Aika semmoisia varmoja ja koeteltuja ratkaisuja haetaan enemmän kuin innovatiivisia pikavoittoja. Ihan johtuen toiminnan luonteesta.”*

H9: *”Säännösten noudattaminen on tietenkin tärkeämpää, että ei voi olla liian innovatiivinen, että sääntöjen, lakien ja asetusten mukaan pitää mennä.”*

Nämä lainaukset selittävät taustoittavan kyselyn tulosta säännöstenmukaisuuden ja innovatiivisuuden tärkeydestä. Koska kyselyn tulosten mukaan datan hyödyntäminen ja joustavuus koettiin tärkeäksi, olisi voinut olettaa, että myös innovatiivisuus nousee säännöstenmukaisuutta tärkeämmäksi asiaksi. Haastatteluiden perusteella voidaan kuitenkin tunnistaa, että toimialan vaatimukset ja säännökset rajoittavat mahdollisuuksia olla innovatiivinen, jolloin sitä ei nähdä niin selkeästi merkityksellisessä osassa kuin esimerkiksi joustavuutta. Tässä tulee kuitenkin huomioida, että haastatteluissa kävi ilmi, että vastaajat tulkitsevat kyselyssä mainitun säännöstenmukaisuuden monella tavalla. Osa tulkitse sen tarkoittavan esimerkiksi datan muotoilua tiettyjen sääntöjen mukaan ja osa liitti asian esimerkiksi toimialasäännöksiin.

Teema ”Yrityksen toiminnan laajuus” muodostettiin yhdistämällä mainintoja datan ja henkilöstön määrän kasvun vaikutuksista toimintamalleihin. Haastateltavat toivat ilmi, että yhtiön koon kasvaessa on yhä tärkeämpää, että toimintamallit ja vastuurakenteet on määriteltä (H6, H10).

H10: *”Organisaation koko on meillä kasvanut, niin kyllähän se tuo väkisin sitä, että pitää tuoda enemmän niitä rakenteita siihen hallintaan ja tekemiseen... aikaisemmin se melkein riitti (ICT-kehityksessä), että käytävällä bongasi jonkun ja sanoi, että tehdään tämänöinen pois. Pari päivää meni niin se oli toteutettu, kun taas nyt se menee sinne tiettyihin rakenteisiin ja sen kautta hoituu. Mutta kun on isompi organisaatio niin se ei vaan toimi enää sillä, että mä huikkaisin sulle, että hei tehdään tommoinen juttu ja sä vastaat no joo, itse asiassa mä voisin sen tällä viikolla jopa tehdä. Koska siellä voi olla kahdeksankymmentäneljä muuta huikkaajaa ja sitten sulla on heti se ongelma, että minkä teet. Että silloin jo täytyy laittaa johonkin listalle ja ruveta priorisoimaan.”*

H10: *”Meillä on paljon enemmän käyttäjiä ja meillä on paljon enemmän dataa ja sovelluksia myös, että se ei ole ihan sama kuin mitä silloin joskus 20 vuotta sitten, että oli yksi sovellus, jota käytti ehkä viisi henkeä, niin se oli helpommin hallittavissa.”*

Teema ”Päätöksentekokulttuuri” muodostettiin yhdistämällä mainintoja suhtautumisesta byrokraatiaan ja päätöksentekorakenteisiin. Haastateltavat eivät halunneet, että datanhallintamalli rajoittaisi tai hidastaisi kokeilemista ja kehitystä (H1, H3, H4).

H4: *”Toivon, että edelleen pystyisi joustavasti ja helposti tekemään uusia juttuja, että ei tulisi ainakaan mitään raateja tai foorumeja tai ryhmiä, joita muodostettaisiin ihmettelemään jotain uusia tarpeita. Tässä yhtiössä tehdään vaan paljon semmoista, että perustetaan joku konklaavi miettimään asioita ja sit se hidastaa kaikkea tekemistä tosi paljon... Jos joku hallintamalli toteutuu ja sitä ruvetaan käytännössä tekemään niin se olisi semmoinen, että se ei ainakaan haittaisi kehittämistä.”*

Hallintamallia ei haluttu myöskään tehdä pelkän hallintamallin vuoksi, vaan nähtiin, että liiketoiminnan tulisi hyötyä siitä. (H2, H8, H10)

H8: *”Ei sitä kannata sen datanhallintamallin itsensä takia rummuttaa.”*

H2: *”Datanhallinta ei saa olla itsetarkoitus vaan sen tulee tukea liiketoimintaa.”*

Haastatteluaineiston nostoista nähdään, että kohdeorganisaation toimintamalleihin vaikuttaa vahvasti yhtiön toimintaympäristö ja toimiala. Organisaation täytyy pystyä olemaan joustava ja heillä täytyy olla kyky reagoida nopeasti muutoksiin, mutta toisaalta turvallisuus ja säännösten noudattaminen on yrityksen toiminnan luonteen kannalta oleellista. Myös yrityksen toiminnan kasvu vaatii raamien ja toimintatapojen määrittämistä, mutta yhtiössä kuitenkin arvostetaan joustavaa ja tehokasta päätöksentekoa, eikä turhaa byrokraatiaa haluta. Haastatteluiden perusteella yhtiössä on tähän mennessä pystytty luovimaan näiden hieman ehkä ristiriitaisten vaatimusten kanssa, sillä haastateltavat olivat yleisesti ottaen melko tyytyväisiä organisaation joustavuuden tasoon ainakin datan hyödyntämiseen ja hallintaan liittyen.

4.2.3 Datanhallinnan rakenteet

Teemat ”Datanhallinnan roolit ja vastuut” sekä ”Määritellyt toimintatavat” yhdistettiin pääteeman ”Datanhallinnan rakenteet” alle. Taulukossa 13 on esitelty teemojen muodostuminen.

Taulukko 13. Teeman ”Datanhallinnan rakenteet” muodostuminen

Pääteema	Teemat	Koodit	Mainittu haastatteluissa
Datanhallinnan rakenteet	Datanhallinnan roolit ja vastuut	Epäselvät roolit	H1, H2, H3, H7, H8, H10
		Vastuuaukko	H2, H3, H4, H5, H9
	Määritellyt toimintatavat	Yhteiset toimintatavat	H2, H3, H5
		Raporttitarpeiden hallinta	H9, H10
		Seuranta ja ylläpito	H3, H4, H9

Aineistosta tunnistettiin mainintoja roolien ja vastuiden epäselvyydestä sekä puutteista vastuiden määrittelyssä. Näistä maininnoista luotiin teema ”Datanhallinnan roolit ja vastuut”. Haastateltavat toivat esiin, että datanhallinnan roolien vastuut voivat olla epäselviä rooleihin nimetyille henkilöille (H1, H2, H10), jolloin asioita voi jäädä hoitamatta (H10).

H1: *”Tässä on varmaan aika paljon rooleja, jotka on epäselviä vielä heille, joille se rooli tavallaan kuuluisi.”*

H10: *”Silloinhan tehtiin näitä erinäköisiä kuvauksia ja nimettiin eri henkilöitä erilaisiin rooleihin. Mulle itsellä jäi kuitenkin vähän se kuva, että ei ne henkilöt siinä roolissa joko sitten pystyneet tekemään niitä asioita mitä ehkä oli ajateltu tai sitten ne ei välttämättä edes ihan tienneet, että mitä mun nyt sitten ihan oikeasti pitäisi tässä tehdä.”*

H10: *”Me on tehty esimerkiksi joitain raporteja ja yhtäkkiä se raportti ei sitten enää toimiikaan... ..joitain raporteja tarvitaan muutaman kerran vuodessa, niin sitten se on melkein niin, että edellispäivänä rupeaa niitä dataa sieltä hakemaan ja sitten kun se (raportti) ei toimi niin se on aika ikävää. Siinä mielessä just tää tämmöinen datanhallinta ja voisiko sanoa datan huolto on oleellista, että siinä on selkeät rakenteet ja vastuut ja toimintamallit.”*

Nämä maininnat ovat linjassa taustoittavan kyselyn vastausten kanssa. Vastauksista kävi ilmi, että vain yhdelle vastaajalle hänen roolinsa datanhallintaan liittyen oli täysin selkeä.

Roolien omistajien tietoisuuteen liittyvien haasteiden lisäksi haastateltavat toivat esiin, että vastuurakenteet ovat välillä myös muulle organisaatiolle epäselviä ja selkeämmistä roolituksista hyödyttäisiin (H3, H7, H8, H10).

H7: *”Ehkä kaipaisi vähän semmoista selkiytymistä siihen kokonaisuuteen, että kuka vastaa siitä (datanhallinnan) strategisesta kehittämisestä ja millä tavalla se operatiivinen kehittäminen siihen linkittyy ja mistä saa mihinkin apua. On se kyllä mennyt parempaan suuntaan niin kun tosi paljon tässä vuosien varrella, mutta jotenkin kaipaisi semmoisen yhden kuvan, josta näkisi vaikka, että mistä kukin vastaa ja mihin niitä asioita voidaan viedä.”*

H8: *”Varsinkin, jos meillä olisi paremmin määritellyt roolit, jossa olisi muutakin kuin yksi nimi joka roolissa, niin olisi helpompi esimerkiksi niihin projekteihin laittaa aina oikea vastuuhenkilö mukaan. Ja jos organisaatio tietää, että tämä henkilö vastaa nyt esimerkiksi sähköasematiedoista, niin sitten sellaiseen projektiin, jossa tarvitaan niitä tietoja, niin osattaisiin oletuksena jo ottaa se tietovastaava mukaan.”*

Haastateltavat tunnistivat myös, että kohdeorganisaatiossa on asioita, joista ei heidän mielestään vastaa selkeästi kukaan (H2, H3, H4, H5, H9). Osa näistä maininnoista liittyi asioihin, jotka ovat monimutkaisia ja niiden hallitsemiseen tarvitaan useamman tietoa-alueen tai yksikön panosta, joten niiden vastuuttaminen on sen vuoksi hankalaa (H4, H5).

H3: *"Ihan kaikelle datalle ei ole nimettyä vastuuhenkilöä. Yksi mittaustieto oli väärin, kun siellä oli laskenta jotenkin hajonnut. Niin siinäkin olisi ollut aika paljon apua, että olisi ollut joku yksi henkilö, joka olisi oikeasti tuntenut, että mikä se on ja mitä siellä oikeasti laske- taan ja kuka on tehnyt siihen muutoksia viime aikoina."*

H5: *"Niin pitkään, kun se on täällä meidän omalla tietoa-alueella niin ok, tää on meidän vastuu. Me otetaan tää ja kyllä tää tästä, juna kulkee. Sitten on semmoista vähän ei kenenkään maata ja siihen liittyviä tietoja, jotka... ..ei oikein ihan suoraan kuulu kenel- lekään. Toisaalta taas ei sitä nyt kannata ehkä liioitellakaan"*

Edellä avatuista haastattelumaininnoista ja taustoittavan kyselyn tuloksista nähdään, että kohdeorganisaatiossa kaivataan roolien ja vastuiden selkeyttämistä. Yhdeksän haastateltavaa kymmenestä nosti vähintäänkin jollain tasolla esiin haasteita roolitukseen ja vastuihin liittyen. Kaikista merkityksellisimmäksi haasteeksi roolitukset ja vastuut ko- ettiin asiantuntija- ja tiiminvetäjätasolla. Kohdeorganisaation ylin johto oli tietoinen, että vastuuhaasteita oli noussut välillä esiin, mutta ne eivät näkyneet yhtä merkittävästi ylim- män johdon edustajalle. Tätä voi selittää se, että epäselvien vastuiden vaikutuksia tun- nistettiin lähinnä käytännön toiminnan tasolla: asioita voi jäädä tekemättä tai tekemisen edistäminen on hitaampaa, koska ei ole selkeää, keneltä tukea voi pyytää.

Haastatteluiden perusteella vaikuttaa siltä, että organisaatiossa on panostettu vastuura- kenteiden luomiseen ja henkilöiden nimeämiseen rooleihin, mutta rakenteiden jalkau- tusta käytäntöön ei ole suunniteltu tarpeeksi. Jos vastuuhenkilö ei tiedä asian olevan hänen vastuullaan, ei asia tule hoidetuksi. Tämä voi myös näyttäytyä muulle organisa- tiolle siltä, että asia ei ole kenenkään vastuulla, vaikka periaatteessa paperilla vastuu olisi jo jollekin määrätty. Vastuurakenteiden jalkautukseen liittyen haastatteluaineistosta voidaan myös tunnistaa tarve vastuurakenteiden näkyvyyden parantamiselle. Tiedon tu- lisi olla helposti löydettävissä silloin, kun sitä tarvitaan.

Aineistosta tunnistettiin mainintoja yhteisten toimintatapojen puuttumisesta, raporttitar- peiden hallinnan puutteesta ja datan laadun seurannan puutteesta. Näistä maininnoista luotiin teema "Määritellyt toimintatavat". Haastateltavat toivat ilmi, että tietoa-alueet tekevät useimmiten omia asioitaan erillään (H3, H5) ja yhteisten linjausten tekeminen on vaikeaa (H2).

H2: *"Meillä ei ole mitään yhtenäisiä linjauksia (sovellus- ja datakehitykseen). Me ei pystytä tekemään semmoisia, koska me ollaan jotenkin huonoja siinä. Me ei pystytä luomaan sellaisia yhteisiä käytäntöjä sellaisista asioista, joissa pitää tehdä vaikeita päätöksiä eli valita kahden tavan väliltä toinen vaikkakin molemmissa on hyvät ja huonot puolet. Me ei pystytä valitsemaan vaan se jää ilmaan se asia ja sitten osa tekee toisella ja osa toisella tavalla."*

Tätä löytöä tukee myös taustoittavan kyselyn tulokset, joiden mukaan puolet vastaajista koki, että eri tiimit eivät tee tarpeeksi yhteistyötä datan hallinnan ja käytön suhteen. Tiimien ja toimintojen yhteistyötä kuitenkin tarvitaan, jotta yhteisiä toimintatapoja voidaan luoda. Yhteisten toimintatapojen ja yhteistyön vähyyttä voi osin selittää luvussa 4.2.1 nostettu haastateltavan 5 mainitsema kohdeorganisaation taipumus keskittyä asioissa eteenpäin menemiseen raamittamisen sijaan. Käytännön tekemisen vahva priorisointi voi toisaalta olla merkki siitä, että yhteisten toimintatapojen ja rakenteiden, kuten datan hallinnan, hyödyt organisaatiossa eivät ole täysin selviä, jos käytännön tekeminen kiillaa usein niiden kehittämisen edelle.

Haastateltavat myös kokivat, että raporttitarpeita ja kehitystä tulisi hallita paremmin, sillä raportteja on tehty melko pienellä kynnyksellä, mikä on johtanut siihen, että joidenkin raporttien käyttöaste on pieni (H9, H10).

H9: *"Pitäisi pitää jonkun näköinen kontrolli kuitenkin siellä, että ei sekään sitten oikeasti kyllä toimi, että jokainen pyytää oman raportin."*

H10: *"...Sitten niitä makaa tuolla pilvin pimein ja niiden käyttöaste ei välttämättä ole kauden suuri ja ne voi vähän unohtuakin sinne. Kyllä siinäkin (raportoinnissa) pitäisi aina miettiä sitä, että mitkä on ne oleelliset raportit, joita ihan oikeasti sitten seurataan ja joista on oikeasti sille liiketoiminnalle hyötyä, eikä niin, että ne on vähän semmoista menovettä, että joskus voi vilkaista, kun on sadepäivä."*

Haastateltavat nostivat esiin myös seurannan ja ylläpidon puuttumisen datan laatuun ja raporttien toimivuuteen liittyen. Nyt organisaatiossa ei aina välttämättä huomata, jos jokin data ei päivity (H3, H4) tai raportilla on virheitä (H9).

H3: *"Datan laadun seuranta olisi ehkä semmoinen, mitä tarvittaisiin enemmän. Sellaista ei ainakaan meillä ole juurikaan mitenkään automatisoitu. Huomattaisiin, että jos meillä vaikka joku data ei päivity."*

H9: *"Se ylläpitokin on tosi tärkeä. Ei pelkästään se, että saadaan luotua (raportteja), vaan saadaan myös pidettyä ne ajan tasalla ja toimivana, kun siellä taustalla tapahtuu koko ajan kuitenkin jotain pientä eloa ja muutosta, niin pystyy sitten päivittämään niitä."*

Teemaan ”Määritellyt toimintatavat” liittyvien haastattelumainintojen perusteella voidaan todeta, että organisaatiossa keskitytään paljon asioiden edistämiseen, mutta toimintatapojen määrittäminen ja asioiden hallinta saattaa välillä jäädä taka-alalle.

4.2.4 Datanhallinnan toteutus

Teemat ”Datanhallinnan jalkauttaminen”, ”Yhteistyö”, ”Aikaresurssit”, ”Osaaminen” ja ”Teknologiat” yhdistettiin pääteemaksi ”Datanhallinnan toteutus”. Teemojen muodostuminen on avattu taulukossa 14.

Taulukko 14. Teeman ”Datanhallinnan toteutus” muodostuminen

Pääteema	Teemat	Koodit	Mainittu haastatteluissa
Datanhallinnan toteutus	Datanhallinnan jalkauttaminen	Käytäntöjen jalkauttaminen	H1, H2, H4
		Tietoisuus datanhallinnan rakenteista	H1, H2, H3, H4, H7, H10
	Yhteistyö	Yhteistyö dataosaajien ja liiketoiminnan välillä	H1, H2, H6, H7, H9
		Tiimien välinen yhteistyö	H1, H2, H3, H4, H5, H6, H7, H10
	Aikaresurssit	Datanhallinnan roolien aikaresurssit	H2, H5, H8
		Liiketoiminnan aikaresurssit	H1, H2, H7, H6, H10
	Osaaminen	Osaaminen datanhallinnasta	H7
		Datakyvykkyydet	H1, H3, H4, H6, H9
	Teknologiat	Tietoallas mahdollistajana	H1, H2, H3, H5, H8, H10
		Teknologioiden jatkuvuus	H10

Haastatteluaineistosta tunnistettiin mainintoja datanhallinnan periaatteiden käytännön toteutuksesta sekä puutteista organisaation tietoisuudessa datanhallinnan rakenteista ja periaatteista. Näistä maininnoista muodostettiin teema ”Datanhallinnan jalkauttaminen”. Haastateltavat toivat esiin, että se, miten datanhallinnan periaatteet näkyvät käytännössä ja ylätasolla suunniteltu asia jalkautuu käytännön toimintaan, ei ole aina selkeää tai loppuun asti mietittyä (H1, H2, H4).

H1: ”...se (datanhallinta) on hyvin varmasti ajateltu ja suunniteltu, mutta sitten jalkautus käytäntöön niin se on toinen asia. Ehkä siihen kaipaisi konkretiaa... konkreettinen esimerkki on se, että viralliselle raportointikerrokselle julkaistulle datalle pitäisi antaa laatus lupaus, niin miten se käytännössä toteutuu.”

H4: ”Se menee semmoiseksi konsulttikielen tosi nopeasti, että esitellään tällaisia suunnitelmia ja road mappeja. Ja miten ne sitten käytännössä toteutuu niin ei ole tässä muutamana vuoden sisällä mielestäni mitään muuta konkreettista muutosta tapahtunut.”

Haastateltavat mainitsivat myös, että tietoisuus datanhallinnasta ja sen politiikoista tai rakenteista ei saavuta kaikkia (H1, H2, H3, H4, 7, H10).

H2: *"Ei meidän yksikössä tiedetä yhtään mitään mistään datanhallinta poliitikoista tai käytänteistä... ...Kun pitää joku ongelma ratkaista, joka vaatii datan kehitystä, niin ei sitten oikein tiedetä, että mistä lähteä kyselemään ja kenen kautta ja kenen kuuluisi tehdä, kun itse ei osata eikä ehditä tehdä. Niin sitten sitä kysellään sieltä ja täältä ja ehkä se lopulta päättyy oikeaan kohteeseen."*

H3: *"Suurin osa fingridläisistä ei tiedä mitä on datanhallinta tai mikä meidän datanhallintamalli on."*

H4: *"Mä oon aika monesta raportista itse vastuussa ja mulla ei ole todellakaan selvää se, että kuka tuolla hallinnasta vastaa ja mitä siellä tapahtuu. Mun mielestä se kanssa kertoo jo jotain siitä, että mä teen aika paljon töitä noiden parissa, mutta ei ole siitä mitään hajua."*

Haastateltavat nostivat kuitenkin esiin, että kaikki organisaation työntekijät eivät myöskään tarvitse kaikkea tietoa datanhallinnasta (H10) ja viestintä ei saisi kuormittaa liikaa (H7).

H10: *"Aika ajoin jonkinlainen kertaus niihin kokonaisuuksiin voisi olla jopa ihan hyvä. Tietysti pitää muistaa, että jos me puhutaan datanhallinnasta ja sen hyödyntämisestä, niin eihän meillä kaikki ole semmoisia henkilöitä, jotka määrittelisi raporteja tai muuta. Ei meillä ole 550:tä semmoista henkilöä. Mutta meillä voi olla 550 semmoista henkilöä, jotka niitä sitten käyttää, jolloin niitten ei tarvitse oikeastaan tietää muuta kuin se, että tällainen raportti on ja tätä voi käyttää."*

H7: *"Tietenkin sitten aina on kiire tai semmoinen kiireen tuntu, että kaikkea pitää tehdä, ja sitten jos tulee lisää palaveriteita niin se olisi tärkeää, että ei kuitenkaan liikaa kuormita liiketoimintaa."*

Vastauksista nähdään, että kohdeorganisaatiossa olisi tarve parantaa datanhallinnan periaatteiden ja käytäntöjen jalkautusta. Tämä on linjassa edellisen alaluvun löydösten kanssa roolitukseen liittyen. Ei ole riittävää, että periaatteet ja rakenteet suunnitellaan ylätasolla ja paperilla, vaan niiden jalkauttaminen käytäntöön täytyy myös suunnitella ja varmistaa. Datanhallinnan periaatteiden toteuttamisesta vastaavien henkilöiden tulisi ymmärtää, mitä periaatteilla tavoitellaan ja mitä vaatimuksia ne käytännössä asettavat ennen kuin he voivat varmistaa, että niitä noudatetaan.

On myös huolehdittava, että organisaatiossa ollaan tarvittavilta osin tietoisia näistä periaatteista ja käytänteistä sekä niihin liittyvistä muutoksista. Koko henkilöstö ei kuitenkaan

tarvitse kaikkea tietoa. Haastatteluissa nousi selkeästi esiin, että kohdeorganisaation henkilöstöllä on erilaisia tarpeita datanhallintaan liittyvälle viestinnälle ja ne tulisi ottaa huomioon. Monet haastateltavat nostivat esiin, että organisaatiossa on paljon henkilöstöä, jotka eivät tiedä datanhallinnasta mitään. Tämä ei kuitenkaan ole yksiselitteinen osoitus viestinnän epäonnistumisesta. Kaikkien henkilöiden ei välttämättä tarvitse tietää datanhallinnan määritelmää, jos he kuitenkin ovat tietoisia heitä koskevista datanhallinnan käytänteistä. Viestinnässä voidaan kuitenkin tunnistaa kehityskohteita, sillä haastatteluissa nousi esiin, että myös ihmiset, jotka tunnistavat itsellään dataan ja sen hallintaan liittyvän roolin, eivät aina ole tietoisia datanhallinnan käytäntöjen ja periaatteiden muutoksista.

Taustoittavan kyselyn tulokset tukevat tätä pohdintaa. Kolme vastaajaa oli osittain eri mieltä siitä, että tieto dataan ja sen hallintaan liittyvien käytänteiden ja menettelytapojen muutoksista saavuttaa heidät hyvin. Toisaalta kuitenkin kuusi vastaajaa kymmenestä oli asiasta täysin tai osittain samaa mieltä. Vastauksissa ei ole selkeää jakautumista organisaatiotasojen välillä. Organisaatiotasoja enemmän eroja näyttäisi selittävän se, onko vastaaja kuinka tiivisti datakehityksessä mukana. Viestinnässä kehitettävää näkivät enemmän henkilöt, jotka eivät olleet enää niin tiiviisti mukana datakehityksessä sekä henkilöt, jotka edustivat datakehityksessä enemmän liiketoiminnan puolta. Tämän perusteella datakehityksen ulkokehän sidosryhmille kohdistuvaa viestintää olisi tarvetta kehittää enemmän kuin itse datakehityksestä vastaaville henkilöille kohdistuvaa viestintää.

Aineistosta tunnistettiin myös mainintoja liiketoiminnan ja datakehityksen välisestä yhteistyöstä sekä tiimien ja toimintojen välisestä yhteistyöstä. Näistä muodostettiin teema ”Yhteistyö”. Liiketoiminnan edustajien näkökulmasta yhteistyö datakehityksen kanssa on ollut sujuvaa (H6, H7). Toisaalta haastatteluissa nousi esiin myös se, että kaikilla tiimeillä ei ole yhtä suoraa yhteyttä datakehitykseen, jolloin he välillä tuntevat jäävänsä tarpeidensa kanssa hieman yksin (H2). Datakehityksen näkökulmasta haastatteluissa nousi esiin toive liiketoiminnan selkeämmästä ja proaktiivisemmasta roolista muun muassa raporttikehityksessä (H1).

H1: ”Liiketoiminnalta tarve tulee yleensä aika myöhään. Eli havahtuvat siihen, että jollekin raportille olisi nyt käyttöä, niin se tulee yleensä tosi myöhäisessä vaiheessa ja pitää sitten kiireellä lähteä tekemään.”

Myös datakehitystiimien välisessä yhteistyössä nousi esiin kehityskohteita (H1, H2, H4, H5, H10) ja haastateltavat kokivat, että organisaatiossa voitaisiin myös yleisellä tasolla tehdä enemmän tiimien välistä yhteistyötä tehokkuuden lisäämiseksi (H3, H7).

H1: *"Ei me mun mielestä hirveästi ole eri tiimien välillä kommunikoitu, että miten he tekee asioita ja miten me tehdään asioita... ..Esimerkkinä vanhasta tietovarastosta irtaantuminen. Ei mulla ole mitään käryä tällä hetkellä, että missä vaiheessa muut menee sen suhteen. Meillä on kova kiire saada oma tontti siivottua, mutta ei mitään käsitystä, että kuinka realistista muilla domain tiimeillä on sen suhteen."*

H5: *"Varmasti olisi tarvetta niiden hyvien käytänteiden ja ahaa elämysten jakamiseksi. Tämmöiset meinaa nyt vähän jäädä tässä kaiken muun kiireen ohella."*

Yhteistyön puutteiden nähtiin johtavan muun muassa päällekkäisen työn tekemiseen, sillä välillä tiimit ovat saattaneet työstää jotain asiaa tietämättä, että toisessa tiimissä samaa asiaa on jo työstetty (H2). Toisaalta kuitenkin ylin johto (H6) nosti esiin, että isossa kuvassa yhteistyö organisaation sisällä toimii hyvin.

H6: *"No me lähtökohtaisesti ei voida olla kauhea siilo-organisaatio. Totta kai kaikilla on se oma vastuualueensa... .. Mutta kyllä me nyt kuitenkin sitä yhteistyötä tehdään ja nyt tietysti tämän uusitun strategian kautta, niin se oikeastaan on sitten tuotu vielä vähän näkyvämmäksi."*

Taustoittavan kyselyn tulokset kuitenkin tukevat näkemystä siitä, että yhteistyötä voitaisiin tehdä vielä enemmän ainakin datan hallintaan liittyen. Puolet vastaajista koki, että kohdeorganisaatiossa ei tehdä tarpeeksi yhteistyötä datan hallinnan ja käytön suhteen. Kaikki asiantuntijat kuuluivat tähän ryhmään. Tästä voidaan tulkita, että yhteistyöhön liittyvät ongelmat näkyvät erityisesti asiantuntijatasolla, mutta ne eivät hidasta tai estä toimintaa niin merkittävästi, että asia nousisi esiin merkittävässä määrin johtotasolla asti. Yhteistyötä kehittämällä voitaisiin kuitenkin tehostaa toimintaa ja vähentää asiantuntijoiden räsitusta.

Aineistosta tunnistettiin myös mainintoja työajan resursointiin liittyen. Haastateltavat toivat esiin, että datanhallintamallissa määriteltyjen roolien omistajilla ei ole tarpeeksi aikaa hoitaa roolin vastuulla olevia tehtäviä (H1, H2, H5, H8).

H2: *"Siellä on määritelty ne roolit, mutta ei ne henkilöt oikeasti tee niitä hommia, koska ne on kädet täynnä niitä substanssiahommia. Ja sitten tää on tällainen sivu mikä niille on annettu, että sä vastaat nyt näistä, niin ne ei oikeasti ehdi tekemään niitä. Noi hommat vaatii yllättävän paljon aikaa, jos niitä haluaa tehdä kunnolla, mutta sitä ei aina kaikkialla tässä talossa tunnisteta, että siellä pitää tehdä aikaa."*

Lisäksi liiketoiminnalla ei nähty aina olevan tarpeeksi aikaa osallistua datakehitykseen (H1, H2, H7, H10).

H1: *”Toivoisin liiketoiminnalta resurssia, että sitä pystyisi kiinnittämään selkeämmin siihen datatekemiseen.”*

H10: *”Siinä on vähän just se haaste myös, että jos se tarve tulee sieltä liiketoiminnasta, niin liiketoiminnalla tietysti pitäisi olla sitten myös aikaa sekä testata ja määritellä se raportti... ..varmaan olisi enemmän myös ajatuksia ideoita, mutta niitä ei vaan sitten ajan vuoksi pystytä kaikkia toteuttamaan, koska se vaatii just nimenomaan sitä liiketoiminnan osaamista ja tietämystä, että miten se raportti toteutetaan.”*

Toisaalta haastatteluissa tuotiin kuitenkin ilmi, että ajan niukkuus ei ole vain huono asia, sillä se pakottaa keskittymään olennaiseen (H10). Resurssien puute ei myöskään ole organisaatiossa niin näkyvä ongelma, että asiaa olisi tuotu merkittävässä määrin ilmi ylimmälle johdolle asti (H6). Näistä ajan resursointiin liittyvistä maininnoista luotiin teema ”Aikaresurssit”.

Taustoittavan kyselyn tulokset tukevat haastatteluiden mainintoja resurssien niukkuudesta. Kuusi kymmenestä vastaajasta koki, että datanhallintaan osallistuvilla henkilöillä ei ole tarpeeksi aikaa tehtäviensä toteuttamiseen. Yksikään vastaaja ei ollut asiasta samaa mieltä. Tämä tukee aiemmin todettua näkemystä siitä, että organisaatiossa ei ole huomioitu tarpeeksi datanhallinnan jalkautusta. Jotta datanhallintamallissa määriteltyjen roolien vastuut eivät jäisi vain kirjauksiksi paperille, täytyy organisaation huolehtia, että roolien omistajilla on tarpeeksi aikaa tehtäviensä toteuttamiseen.

Teema ”Osaaminen” muodostettiin maininnoista, jotka liittyivät datanhallinnan osaamiseen sekä organisaation henkilöstön datakyykykkyyksiin. Yksi vastaaja toi ilmi, että haluaisi parantaa omaa osaamistaan datanhallintaan liittyen, mutta uskoi kyllä muualla organisaatiossa olevan siitä tarpeeksi osaamista (H7). Liiketoiminnasta löytyi myös halukkuutta opetella itse tekemään data-analyyseja, jos työkalut, kuten tekoäly, tekee siitä tarpeeksi ketterää ja yksinkertaista (H9). Vastaajat näkivät toisaalta myös muualla organisaatiossa mahdollisuuksia kehittää dataosaamista. Kehitysmahdollisuuksia nähtiin liiketoiminnassa ja heidän kyyvyssään määritellä omia tarpeitaan (H1) sekä työntekijöiden kyyvyssä tehdä asioita uudella tavalla, jotta datasta voitaisiin ottaa kaikki irti (H3). Organisaatiossa nähtiin kuitenkin olevan todella päteviä dataosaajia, jotka pystyvät toteuttamaan melkein mitä tahansa (H6).

Taustoittavan kyselyn tuloksista voidaan kuitenkin päätellä, että kohdeorganisaatio voisi vielä kehittää datanhallintaan liittyvää osaamista. Vastaukset väitteeseen ”Koen, että organisaatiossa on tarpeeksi ymmärrystä ja osaamista ja datanhallintaan liittyen” vaihtelivat paljon. Puolet vastaajista oli kuitenkin väitteen kanssa osittain tai täysin eri mieltä.

Koska aiemmin todettiin etäisyyttä liiketoiminnan ja datanhallinnan välillä ja haastateltavat kokivat liiketoiminnan tietoisuuden datanhallinnasta vähäiseksi, voidaan olettaa, että datanhallinnan osaamisen kehittämistarpeet nähdään pääasiassa juuri liiketoiminnan puolella. Osaamisen kehittämisellä voitaisiin tehostaa datakehitystä sekä lisätä työntekijöiden uskallusta ja halua tehdä asioita uusilla tavoilla.

Haastateltavat toivat myös esiin teknologian merkityksen datan hallinnassa ja hyödyntämisessä. Monet näkivät uuden tietoallasratkaisun tuovan uusia mahdollisuuksia ja tehostavan toimintaa (H1, H2, H3, H5, H10). Yleisesti ottaen kohdeorganisaation teknologiat ja työkalut koettiin toimiviksi, organisaation koettiin vain tarvitsevan aikaa ja kykyä hyödyntää niitä (H8).

H2: ”Kyllä meillä aika hyvät työkalut ja modernit teknologiat on käytössä. Nyt on pistetty tietoallas tulille ja on kaikennäköiset vehkeet. Niin me ei olla kauhean rajoitettuja nyky-päivänä siihen, että millä työkaluilla tehdään. Mä oon sanonut tässä viime aikoina, että nykyään ei ikinä ole kysymys enää siitä, että voiko jotain tehdä, vaan kysymys on siitä, että millä se kannattaa tehdä.”

Siitä huolimatta, että teknologiat koettiin toimiviksi, haastatteluissa kävi ilmi toive teknologioiden jatkuvuudesta, sillä niiden käyttöönottojen koettiin vaativan paljon työtä.

H10: ”Ehkä se toinen näkökulma siihen on, että olisi tietyllä tavalla semmoista jatkuvuutta, ettei käy niin, että kun me saadaan tämä data lake tästä nyt niin kun uitua rannalta toiselle niin sitten menee kolme vuotta ja täytyisi panna taas kaikki uusiksi. Ne on kuitenkin omalla tavallaan isoja ja raskaita prosesseja ja projektejakin, että siinä mielessä semmoiset kestävätkin ratkaisut on hyviä.”

Näistä maininnoista muodostettiin teema ”Teknologiat”. Tuloksista nähdään, että kohdeorganisaatiossa on panostettu teknologioihin, jotka edistävät datan hallintaa ja hyödyntämistä. Haastatteluiden perusteella teknologiat nähdään organisaatiossa tärkeinä mahdollistajina. Parhaimmatkaan teknologiat eivät kuitenkaan varmista datan tehokasta ja asianmukaista hyödynnettävyyttä, jos hyödyntämiseen ja hallintaan liittyvät toimintatavat ja prosessit eivät ole loppuun asti mietittyjä. Kohdeorganisaatiossa selkeästi ymmärretään tämä, sillä se on työstänyt molempia. Kuitenkin haastatteluiden perusteella organisaatiossa ollaan hieman edellä teknologioissa ja työkaluissa, sillä haastatteluiden perusteella tunnistetut kehitystarpeet liittyvät enemmän rakenteiden ja toimintatapojen jalkauttamiseen.

4.2.5 Datanhallinnan hyödyt

Datanhallinnan hyödyt muodostettiin viimeiseksi pääteemaksi yhdistämään teemat ”Tunnistetut datanhallinnan hyödyt” ja ”Hyötyjen konkretisointi”. Teemojen muodostuminen on kuvattu taulukossa 15.

Taulukko 15. Teeman ”Datanhallinnan hyödyt” muodostuminen

Pääteema	Teemat	Koodit	Mainittu haastateluissa
Datanhallinnan hyödyt	Tunnistetut datanhallinnan hyödyt	Tehokkuus	H2, H4, H8, H9
		Datan hyödynnettävyys	H2, H10
	Hyötyjen konkretisointi	Hyötyjen todentaminen	H2, H4, H6, H7

Aineistosta tunnistettiin mainintoja datanhallinnan tuomista hyödyistä tehokkuuteen ja datan hyödynnettävyyteen liittyen. Näistä maininnoista muodostettiin teema ”Tunnistetut datanhallinnan hyödyt”. Datanhallinnan nähtiin tuovan tehokkuutta selkeiden vastuurakenteiden kautta (H2, H4). Kuitenkin esiin tuli myös skeptisyyttä siitä, miten käytännössä hyötyä tuottavat vastuurakenteet toteutetaan (H4). Datanhallinnan nähtiin myös tukevan IT-hankkeita ja datan hyödyntämistä ja näin lyhentävän esimerkiksi projektien läpimenoaikaa (H8). Lisäksi datanhallinnan nähtiin mahdollistavan datan löydettävyys ja yksien samojen datasettien hyödyntäminen koko organisaatiossa (H9).

H8: *”Jo lähtökohtaisesti tilanne on se, että datasta meidän yksikön toiminta muodostuu. Ja sitten vaan se, miten asioita voidaan tehdä paremmin tai vastata uusiin haasteisiin, niin datanhallinta voi mahdollistaa sen.”*

H2: *”Toiminta voisi olla aika tehokasta, jos olisi selkeämpää ja ihmiset tietäisivät, mitkä on ne roolit ja mitä putkea pitkin mimmoiset asiat pitää lykätä eteenpäin.”*

Lisäksi aineistosta tunnistettiin mainintoja vaikeuksista uskoa tai konkretisoida datanhallinnan hyötyjä ja toimivuutta käytännössä. Näistä maininnoista muodostettiin teema ”Hyötyjen konkretisointi”. Eräs vastaajista koki, että datanhallinnan hyödyt pitäisi pystyä tuomaan näkyviksi käytännön tasolla. Hän näki siinä mahdollisuuden lisätä henkilöstön sitoutumista datanhallintaan (H2).

H2: *”Jotenkin pitäisi esittää sellaisia konkreettisia esimerkkejä, että onnistuttiin tekemään joku ratkaisu tosi nopeasti ja helposti sen takia, että edellisetkin ratkaisut eri asioihin on tehty samalla tavalla ja ne sitten tuki tätä seuraavaa. Jos me pystytään näyttämään esimerkkejä, kuinka nopeasti ja näppärästi saatiin tehtyä jotain ratkaisuja, niin ne innostaa, että hei minäkin haluan, kun kerta tuokin onnistui noin helposti. Jos koetaan, että me ei*

pystytäkertomaan kerta kaikkiaan, että mitä hyötyä tää hallinta tuo, niin sitten voi miettiä, tuoko se mitään hyötyä.”

Osa myös tunnisti, että heillä itselläänkin oli ollut alkuun vaikea hahmottaa datanhallinnan hyötyjä.

H6: ”Mäkään en varmasti heti alkuun edes oikein ymmärtänyt, kuinka tärkeästä asiasta siinä loppu viimein oli kysymys, että mitä oltiin tekemässä. Sittemmin tullut vähän ymmärrystä tässä, kun matka on jatkunut, että olipas hyvä, että jo silloin lähdimme sitä tekemään.”

Taustoitettavan kyselyn mukaan kaikki vastaajat olivat sitä mieltä, että datanhallinta tuo vähintään jossain määrin hyötyjä kohdeorganisaatiolle. Hyödyt olivat täysin selkeitä kolmelle vastaajalle mutta vain osittain selkeitä kuudelle vastaajalle. Nämä vastaukset tukevat haastatteluiden tuloksia. Haastatteluiden tuloksista käy ilmi, että pääosin datanhallinta nähdään hyödyllisenä asiana, mutta konkreettisia hyötyjä kaikki haastateltavat eivät nimenneet. Luvussa 4.2.1 todettiin, että kohdeorganisaatiossa usein keskitytään mieluummin asioiden edistämiseen kuin raamittamiseen. Datanhallinnan kontekstissa tämä saattaa johtua osin liiketoiminnan ja datanhallinnan etäisyydestä, mutta varmasti myös resursseista, sillä kuten aiemmin todettiin, liiketoiminnalle ja datanhallintaan osallistuville henkilöille ei ollut resursoitu tarpeeksi aikaa datakehitykseen ja datanhallintaan osallistumiseen. Kuitenkin, jos datanhallinnan hyödyt olisivat selkeämpiä, työntekijät mahdollisesti priorisoisivat datanhallintaan liittyviä vastuita työlistalla korkeammalle. Olisi varmasti helpompaa saada kiireisiltäkin työntekijöitä työpanosta, jos olisi selkeää, että käyttämällä nyt aikaa esimerkiksi raamitusten ja toimintatapojen miettimiseen, voidaan jatkossa saavuttaa merkittäviä hyötyjä, kuten säästää aikaa ja vaivaa.

5. JOHTOPÄÄTÖKSET

Tässä luvussa esitellään empiirisen tutkimuksen avulla jalostettu teoreettinen viitekehys ja tehdään yhteenveto kohdeorganisaation datanhallinnan vahvuuksista ja heikkouksista.

5.1 Jalostettu teoreettinen viitekehys

Työn kolmannessa luvussa muodostettiin teorian avulla datanhallinnan jatkuvan kehittämisen malli, jonka tarkoituksena on auttaa organisaatioita kehittämään yrityksen toimintaa tukevaa datanhallintaa. Mallia voidaan hyödyntää organisaatioissa, joissa luodaan ensimmäistä datanhallintamallia sekä organisaatioissa, joissa halutaan uudistaa jo käytöön otettua datanhallintaa. Jatkuva parantaminen sisällytettiin malliin PDCA-mallin rakenteen avulla, sillä datanhallinta tunnistettiin jatkuvaksi prosessiksi, jonka tulee mukautua sisäisiin ja ulkoisiin muutoksiin (DAMA 2017; Walsh et al. 2022).

Kehitysmallissa määriteltiin datanhallintamallin luomisen tai uudistamisen kannalta olennaisiksi tekijöiksi datanhallinnan yhteys liiketoimintaan (DAMA 2017; Abraham et al. 2019; Chakravorty 2020; Lis et al. 2023) sekä datanhallintamallin yhteensopivuus ulkoisten ja sisäisten kontekstitekijöiden kanssa (Weber et al. 2009; Otto 2011b; Abraham et al. 2019; Lis et al. 2023). Datanhallintamallin toteuttamisen kannalta oleellisiksi asioiksi määriteltiin johdon tuki (DAMA 2017; Mahanti 2018; Al-Ruithe et al. 2019), työntekijöiden sitouttaminen datanhallintaan (Al-Ruithe et al. 2019; Karkošková 2023), henkilöstön kyvykkyydet (Mahanti 2018; Al-Ruithe et al. 2019) sekä viestintä ja yhteistyö (DAMA 2017; Mahanti 2018; Bento et al. 2022). Lisäksi malliin sisällytettiin datanhallinnan arviointi ja toiminnan vakiinnuttaminen tai muutosten tekeminen. Malli esiteltiin kuvassa 10.

Empiirisen tutkimuksen jälkeen sen tuloksia peilattiin tähän teoreettisessa viitekehyksessä luotuun malliin (kuva 10). Tällä tavalla malli saatiin osittain vahvistettua. Empiirissä tutkimuksessa nousi esiin datanhallintamallin muodostamiseen vaikuttavina tekijöinä liiketoimintalähtöisyys, toimialan säännökset ja vaatimukset sekä päätöksentekokulttuuri. Liiketoimintalähtöisyys tunnistettiin taustalla vaikuttavaksi tekijäksi, joka näkyi erityisesti data- ja analytiikkakehityksessä. Yksi merkittävä syy kohdeorganisaation data- ja analytiikkakehityksen liiketoimintalähtöisyyteen oli kuitenkin datanhallinnan linjaus kehityksen viemisestä lähemmäs liiketoimintayksiköitä. Näin ollen myös datanhallinnan taustalla nähtiin selkeä pyrkimys liiketoimintalähtöisyyteen ja liiketoiminnan tarpeiden ja

datanhallinnan vaatimusten välillä tunnistettiin yhteys. Toimialan säännökset ja vaatimukset sekä päätöksentekokulttuuri tunnistettiin suoraan toimintamalleihin vaikuttaviksi tekijöiksi. Kohdeorganisaation toimialan nähtiin olevan melko säänneltyä, joten sen toimintamallien pitää mahdollistaa säännösten ja vaatimusten noudattaminen. Toisaalta kohdeorganisaatiossa arvostettiin joustavuutta ja karsastettiin turhaa byrokratiaa, joten tähän päätöksentekokulttuurin sopimattomien hallintamallien luominen aiheuttaisi todennäköisesti organisaatiossa vastustusta. Näiden empiiristen löydösten perusteella liiketoimintastrategia, toimiala, lait ja sääntely sekä päätöksentekotyylit ja kulttuuri vahvistettiin datanhallintamalliin vaikuttaviksi tekijöiksi.

Teoreettisesta mallista vahvistettiin empiirisessä tutkimuksessa myös osa datanhallinnan toteuttamisen edellytyksistä. Näitä olivat yhteistyö ja osaaminen. Yhteistyöhön ja osaamiseen liittyen kohdeorganisaatiossa voitiin tunnistaa joitain kehitystarpeita, mutta ne katsottiin toteuttamista tukeviksi tekijöiksi, koska niiden huonolla tilalla nähtiin olevan negatiivisia vaikutuksia datanhallinnan toteuttamiseen ja kehittämiseen ja kehittämällä taas nähtiin olevan positiivisia vaikutuksia. Myös johdon tuki nousi välillisesti esiin empiirisessä tutkimuksessa. Haastatteluissa kävi ilmi, että kohdeorganisaation johto pitää datanhallintaa ja sen kehittämistä tärkeänä. Tämä mahdollistaa datanhallinnan aktiivisen kehittämisen. Jos kohdeorganisaation johto ei pitäisi datanhallintaa tärkeänä, kohdeorganisaatiossa tuskin olisi ollut mahdollista muodostaa viime vuosina uusia datanhallinnan rakenteita. Voidaan siis todeta, että datanhallinnalla on kohdeorganisaatiossa johdon tuki ja se vaikuttaa positiivisesti datanhallinnan kehitykseen. Näiden empiiristen löydösten perusteella teoreettisesta kehitysmallista vahvistettiin datanhallinnan toteuttamisen edellytyksiksi siis johdon tuki, yhteistyö sekä henkilöstön kyvykkyydet. Teoreettisesta mallista vahvistettiin osittain myös datanhallinnan vaikutusten seurannan oleellisuus. Empiirisessä tutkimuksessa kävi ilmi tarve datanhallinnan hyötyjen konkretisoinnille, mitä voidaan tehdä esimerkiksi datanhallinnan vaikutuksia seuraamalla ja mittaamalla.

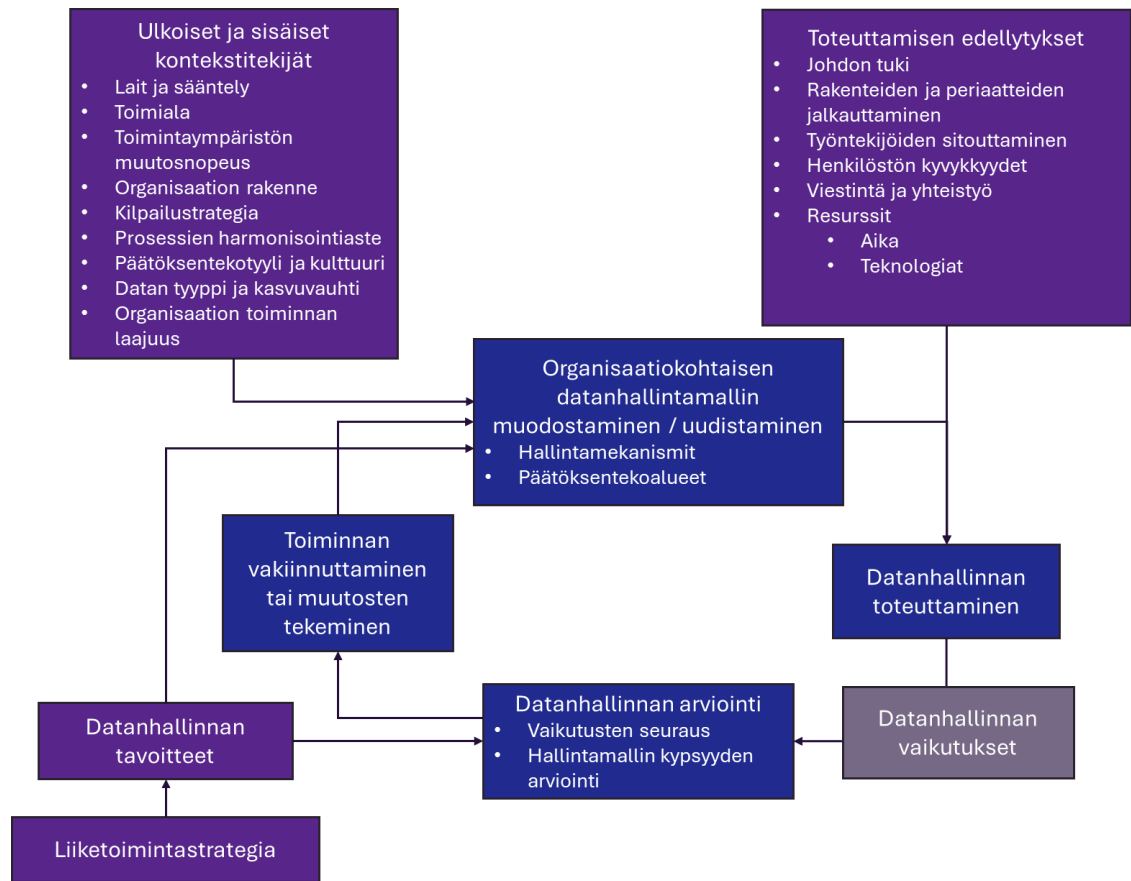
Kuitenkaan kaikkia teoreettisen mallin osia ei pystytty vahvistamaan empiirisen tutkimuksen tulosten avulla. Varsinaisia ristiriitoja mallin ja tulosten välillä ei silti tunnistettu. Lisäksi empiirisessä tutkimuksessa löydettiin uusia datanhallintamalliin vaikuttavia tekijöitä, joita ei aiemmin tunnistettu teoreettisessa viitekehityksessä. Näitä olivat toimintaympäristön nopea muuttuminen ja yrityksen toiminnan laajuus. Teoreettisessa viitekehityksessä tuotiin kyllä ilmi Weberin ja muiden (2009) näkemys organisaation koosta kiisteltynä kontekstitekijänä, mutta kirjallisuuskatsauksessa ei löydetty tutkimuksia, joissa organisaation koko olisi luokiteltu kontekstitekijäksi. Kuitenkin empiirisen tutkimuksen perusteella yrityksen toiminnan laajuus, kattaen henkilöstön määrän ja datan määrän, asettaa vaatimuksia organisaation datanhallinnalle. Uusien datanhallintamalliin vaikuttavien

tekijöiden lisäksi empiirisessä tutkimuksessa tunnistettiin uusia datanhallinnan toteuttamisen edellytyksiä. Näitä olivat datanhallinnan jalkauttaminen ja resurssit, kuten aika ja teknologiat. Teoreettisen viitekehyksen ja empiirisen tutkimuksen tulosten vertailu on koottu taulukkoon 16. Taulukkoon on nostettu muodostetun datanhallinnan kehittämismallin kannalta oleelliset tulokset.

Taulukko 16. *Teoreettisen viitekehyksen ja empiirisen tutkimuksen tulosten vertailu*

	Teoreettinen malli	Empiirinen tutkimus
Yhteys liiketoimintaan	Liiketoimintastrategia	Liiketoimintalähtöisyys
Ulkoiset ja sisäiset kon- tekstitekijät	Lait ja sääntely	Toimialan säännökset ja vaatimukset
	Toimiala	
	Organisaation rakenne	
	Kilpailustrategia	
	Prosessien harmonisointiaste	
	Päätöksentekotyylit ja kulttuuri	Päätöksentekokulttuuri
	Datan tyyppi ja kasvuvauhti	
		Toimintaympäristön nopea muuttuminen
	Yrityksen toiminnan laajuus	
Datanhallinnan toteuttamisen edellytykset	Johdon tuki	Johdon ymmärrys datan ja datanhallinnan tärkeydestä
	Työntekijöiden sitouttaminen datanhallintaan	
	Henkilöstön kyvykkyydet	Osaaminen
	Viestintä ja yhteistyö	Yhteistyö
		Datanhallinnan jalkauttaminen
		Aikaresurssit
		Teknologiat
Datanhallinnan arviointi	Vaikutusten seuranta	Hyötyjen konkretisointi
	Hallintamallin kypsyyden arviointi	

Empiirisen tutkimuksen tulokset yhdistettiin vielä teoreettiseen kehitysmalliin. Malliin tehtiin vain lisäyksiä, sillä kuten edellä todettu, vaikka kaikkia teoreettisen mallin osia ei pystytty vahvistamaan, ei niitä myöskään todettu virheellisiksi. Empiirisen tutkimuksen avulla jalostettu malli datanhallinnan jatkuvasta kehittämisestä organisaation toimintaan sopivaksi on esitelty kuvassa 17.



Kuva 17. *Empiirisen tutkimuksen avulla jalostettu malli datanhallinnan kehittämiseen organisaation toimintaan sopivaksi*

Kuvan 17 mallia voidaan käyttää datanhallinnan luomiseen ja uudistamiseen tietyn organisaation toimintaan sopivaksi. Kirjallisuuskatsauksen perusteella todettiin, että datanhallinta tulisi suunnitella tukemaan liiketoimintastrategiaa ja -tavoitteita, eikä datanhallinnan tavoitteiden tulisi olla irrallisia niistä (Otto 2011b; DAMA 2017; Abraham et al. 2019; Chakravorty 2020; Lis et al. 2023). Tämä sai vahvistusta myös empiirisestä tutkimuksesta. Tämän yhteyden vuoksi kehitysmalli lähtee liiketoimintastrategiasta eli liiketoiminnan tavoitteiden ja tarpeiden ymmärtämisestä. Nämä tavoitteet ja tarpeet taas jalostetaan datanhallinnan tavoitteiksi, jotka vaikuttavat datanhallintamallin muodostamiseen tai uudistamiseen. Mallin muodostamisessa ja uudistamisessa tulee huomioida myös ulkoiset ja sisäiset kontekstitekijät, joita ovat kirjallisuuskatsauksen (Weber et al. 2009; Abraham et al. 2019; Lis et al. 2023) ja empiirisen tutkimuksen perusteella lait ja sääntely, toimiala, toimintaympäristön muutosnopeus, organisaation rakenne, kilpailustrategia, prosessien harmonisointiaste, päätöksentekotyylit ja -kulttuuri, datan tyyppi ja kasvuvauhti sekä organisaation toiminnan laajuus. Kontekstitekijät ovat organisaatiokohtaisia ja niiden huomioiminen mahdollistaa organisaatioille datanhallinnan muokkaamisen omiin ympäristöihin ja tarpeisiin sopivaksi (Weber et al. 2009; Otto 2011b; Abraham et al. 2019; Lis et al. 2023).

Kun nämä taustatekijät on kartoitettu ja huomioitu, voidaan muodostaa datanhallintamalli tai uudistaa jo olemassa olevaa datanhallintamallia. Datanhallintamallin keskeisimmiksi osiksi tunnistettiin hallintamekanismit ja päätöksentekoaalueet. Niiden tulisi olla yhteneväisiä liiketoimintastrategian ja kontekstitekijöiden kanssa. Kirjallisuuskatsauksessa todettiin, että huomiota tulisi kiinnittää erityisesti datastrategian ja liiketoimintastrategian yhteneväisyyteen (Dallemlund & Davenport 2017; Medeiros et al. 2020) sekä kaikkien strategisten tasojen sisällyttämiseen datanhallintaan (Korhonen et al. 2013; DAMA 2017; Chakravorty 2020).

Pelkkä datanhallintamallin muodostaminen ja uudistaminen ei kuitenkaan riitä, sillä se täytyy myös ottaa käyttöön ja sitä tulee toteuttaa. Tämän vuoksi datanhallintamallin muodostamisen jälkeen mallissa on kuvattu datanhallinnan toteuttaminen ja sen edellytykset. Toteuttamisen edellytykset ovat kirjallisuuskatsauksen mukaan elementtejä, joiden puuttuminen voi todennäköisesti johtaa datanhallinnan ja siihen liittyvien projektien epäonnistumiseen (Mahanti 2018). Kirjallisuuskatsauksen ja empiirisen tutkimuksen perusteella toteuttamisen edellytyksiksi tunnistettiin johdon tuki (DAMA 2017; Mahanti 2018; Al-Ruithe et al. 2019), rakenteiden ja periaatteiden jalkauttaminen, työntekijöiden sitouttaminen (Al-Ruithe et al. 2019; Karkošková 2023), henkilöstön kyvykkyydet (Mahanti 2018; Al-Ruithe et al. 2019), viestintä ja yhteistyö (DAMA 2017; Mahanti 2018; Bento et al. 2022) sekä resurssit. Resurssit sisältävät henkilöstön käytettävissä olevan työajan ja organisaation datan hallintaa ja hyödyntämistä tukevat teknologiat. Organisaatioissa tulisi varmistaa, että edellä mainitut edellytykset ovat riittävällä tasolla, tai datanhallinnalla ei välttämättä voida saavuttaa toivottuja vaikutuksia.

Datanhallinnan toteuttamisen jälkeen mallissa on kuvattu vaikutukset, jotka syntyvät datanhallinnan toteuttamisesta. Datanhallinnan vaikutusten voidaan olettaa olevan organisaatiokohtaisia, sillä mallin mukaan organisaatioilla tulisi olla erilaisia datanhallintamalleja, ja datanhallinnan toteuttamisen edellytykset ovat organisaatioissa eri tasolla. Vaikutusten jälkeen mallissa on kuvattu datanhallinnan arviointi. Teoreettisessa viitekehyksessä todettiin, että datanhallinnan onnistumisen arviointi on oleellista, jotta sidosryhmille voidaan viestiä ja näyttää, kuinka hyvin datanhallinta on luonut liiketoiminta-arvoa ja saavuttanut tavoitteita (DAMA 2017). Myös empiirisessä tutkimuksessa tunnistettiin tarve konkretisoida datanhallinnan vaikutuksia. Tämä voidaan tehdä muun muassa arvioimalla ja mittaamalla saavutettuja vaikutuksia. Vaikutusten arvioinnilla voidaan myös tarkistaa, tukeeko malli ja sen toteuttaminen datanhallinnan tavoitteita vai täytyisikö datanhallintaan tehdä muutoksia. Tämän vuoksi mallissa on yhdistetty datanhallinnan tavoitteet nuolella datanhallinnan arviointiin. Datanhallinnan tavoitteet vaikuttavat siihen,

millä tavoin vaikutuksia tulisi seurata. Koska datanhallinnan tavoitteet vaihtelevat organisaatioiden kesken, tulisi datanhallinnan arviointimenetelmien ja mittareiden olla organisaatiokohtaisia. Datanhallinnan vaikutusten arvioinnin lisäksi kirjallisuuskatsauksessa todettiin, että datanhallinnan kypsyyden arviointi kypsyyksmallien avulla voi olla hyödyllistä. Kypsyyksmallien avulla voidaan arvioida aiempien kehitystoimenpiteiden toimivuutta sekä suunnitella uusia kehitystoimenpiteitä (Herselman et al. 2019).

Viimeisenä mallissa on kuvattu toiminnan vakiinnuttaminen tai muutosten tekeminen. Tämän jälkeen palataan taas datanhallintamallin uudistamiseen liiketoiminnan tavoitteiden ja kontekstitekijöiden mukaiseksi. Kehityksmalli on kuvattu ympyrässä, sillä kirjallisuuskatsauksessa todettiin, että datanhallinta on jatkuva prosessi, jonka tulee mukautua sisäisiin ja ulkoisiin muutoksiin (DAMA 2017; Walsh et al. 2022). Se vaatii siis jatkuvaa kehittämistä. Tätä tukee myös kirjallisuudessa tunnistettu pyrkimys huomioida liiketoimintastrategia ja kontekstitekijät mallin suunnittelussa. Näissä molemmissa tapahtuu muutoksia, minkä jälkeen datanhallintamalli ei välttämättä ole enää niiden kanssa yhteensopiva. Tällöin datanhallintamallia tulee tarkastella uudelleen.

Kuten aiemmin todettu kuvan 17 mallia voidaan käyttää datanhallinnan kehittämiseen tietyn organisaation toimintaan sopivaksi. Mallia noudattamalla varmistetaan, että datanhallinnan suunnittelussa otetaan huomioon sellaiset organisaatiokohtaiset ulkoiset ja sisäiset tekijät, jotka vaikuttavat siihen, millainen datanhallintamalli tukee organisaation toimintaa parhaiten. Jos näitä ei huomioida, riskinä on, että muodostetaan datanhallintamalli, joka on ristiriidassa esimerkiksi liiketoiminnan tavoitteiden, toimialan vaatimusten tai säännösten kanssa. Tällöin datanhallinnalla ei saada tuotettua yhtä paljon arvoa tai voidaan jopa aiheuttaa liiketoiminnalle haittaa. Malli varmistaa myös, että datanhallinnan kehittäminen ei pääty datanhallintamallin suunnitteluun vaan myös sen toteuttamisen edellytykset otetaan huomioon. Varmistamalla esimerkiksi johdon tarvittavan tuen, rakenteiden ja periaatteiden jalkauttamisen sekä tarvittavat resurssit, pidetään huolta, että datanhallinnalla on mahdollisuus tuottaa hyötyjä. Mallin muodostamisen PDCA-viitekehyyksen pohjalta varmistaa myös, että datanhallinnan tilaa ja toimivuutta arvioidaan ja kehitetään jatkuvasti, mukauttaen sitä organisaation sisäisiin ja ulkoisiin muutoksiin.

Malli on muodostettu kirjallisuuskatsauksen sekä yhden organisaation tapaustutkimuksen avulla, joten sitä ei välttämättä voida soveltaa kaiken tyyppisissä organisaatioissa. Malli ottaa kuitenkin huomioon organisaatiokohtaiset kontekstitekijät, joten sen käyttö ei ole rajoitettu vain kohdeorganisaation tyyppiin yrityksiin. Mallia soveltaessa tulee kuitenkin huomioida, että erilaisissa organisaatioissa voi olla tarvetta painottaa mallin eri osia ja PDCA-viitekehyyksen mukainen jatkuvan kehityksen malli ei välttämättä sovi kaikille organisaatioille.

5.2 Kohdeorganisaation datanhallinnan vahvuudet ja kehityskohteet

Datanhallinnan kehittämismallin luomisen lisäksi tutkimuksessa haluttiin kartoittaa kohdeorganisaation datanhallinnan nykytilaa. Tutkimuksessa haluttiin selvittää kohdeorganisaation datanhallinnan vahvuudet ja heikkoudet. Datanhallinnan vahvuuksiksi tunnistettiin ymmärrys datan arvosta liiketoiminnalle, ymmärrys datanhallinnan arvosta liiketoiminnalle, johdon tuki, pyrkimys liiketoimintalähtöisyyteen, datanhallinnan rakenteiden aktiivinen kehitys, kontekstitekijöiden kanssa ristiriidaton datanhallintamalli sekä datan hyödyntämistä ja hallintaa tukevat teknologiat (taulukko 17). Vahvuudet koottiin taustoitavan kyselyn ja haastatteluiden tulosten pohjalta.

Taulukko 17. *Keskeisimmät tunnistetut vahvuudet kohdeorganisaation datanhallinnassa*

Vahvuudet	Kuvaus
Ymmärrys datan arvosta liiketoiminnalle	Data nähdään kohdeorganisaatiossa arvokkaana voimavarana kaikilla organisaatiotasolla.
Ymmärrys datanhallinnan arvosta liiketoiminnalle	Datanhallintaa pidetään tärkeänä ja hyödyllisenä asiana.
Johdon tuki	Päälliköt ja organisaation ylin johto ymmärtävät datan ja datanhallinnan merkityksen ja pitävät niitä arvoa tuottavina asioina.
Pyrkimys liiketoimintalähtöisyyteen	Kohdeorganisaatiossa on vahva käsitys siitä, että kehitettävien asioiden pitäisi hyödyttää liiketoimintaa. Data- ja analytiikkakehitys koetaan organisaatiossa hyvin liiketoimintalähtöisenä. Myös datanhallinnan ja liiketoiminnan tarpeiden väliltä voidaan löytää yhteys.
Datanhallinnan rakenteiden aktiivinen kehitys	Vaikka datanhallinnan rakenteiden jalkauttamisessa tunnistettiin haasteita, kohdeorganisaatiossa on tiedostettu rakenteiden tarpeellisuus ja muodostettu erilaisia rooleja ja tietäalajakvoja. Näitä myös kehitetään aktiivisesti.
Kontekstitekijöiden kanssa ristiriidaton datanhallintamalli	Kohdeorganisaation toimintamalleja datan hyödyntämiseen ja hallintaan liittyen pidetään pääosin tarpeeksi joustavana. Ne huomioivat myös toimialan asettamat vaatimukset turvallisuudesta ja säännöstenmukaisuudesta. Toimintamallien kehityksessä on reagoitu yrityksen koon ja datamäärien kasvuun.
Datan hyödyntämistä ja hallintaa tukevat teknologiat	Kohdeorganisaatiolla on käytössä teknologioita, jotka mahdollistavat datan tehokkaan hallinnan ja hyödyntämisen. Niitä myös kehitetään aktiivisesti.

Kuten jo johdannossa todettiin, datasta voidaan luoda arvoa vasta, kun siihen suhtaudutaan tärkeänä resurssina ja sen laadusta ja asianmukaisesta käsittelystä huolehditaan (DeStefano et al. 2016). Näin ollen ymmärrys datan arvosta liiketoiminnalle on kriittinen tekijä organisaation dataresurssien arvon maksimoimisessa, mikä on myös datanhallin-

nan tavoite Otton mukaan (2011a). Jos datan arvoa ei ymmärretä, on hyvin vaikeaa perustella, miksi tarvittaisiin rakenteita ja periaatteita varmistamaan sen säännöstenmukainen ja tehokas käyttö. Kohdeorganisaatiossa tämä ymmärrys on hyvällä tasolla. Kaikki organisaatiotasot olivat yhtä mieltä datan kriittisyydestä liiketoiminnan tavoitteiden saavuttamisen kannalta.

Datan arvon ymmärtäminen luo siis edelleen pohjan datanhallinnan merkityksen ymmärtämiselle. Teoreettisessa viitekehyksessä todettiin, että datanhallinta edellyttää koko henkilöstön osallistumista ja sitoutumista (Weberin et al. 2009; Al-Ruithe et al. 2019). Näiden edellytyksenä on sidosryhmien vakuuttaminen datanhallinnan tuottamasta arvosta (Tsavatewan 2023). Kohdeorganisaatiossa tämä ei kuitenkaan näyttäytynyt haasteena. Suurin osa haastateltavista piti datanhallintaa merkityksellisenä asiana ja koki datanhallinnan luovan arvoa kohdeorganisaatiolle. Tämä käsitys kattoi kaikki organisaatiotasot, mukaan lukien päällikkötason sekä ylimmän johdon. Tämän perusteella voidaan todeta, että datanhallinnan kehityksellä on takanaan myös organisaation johdon tuki. Kirjallisuuskatsauksessa selvisi, että johdon tuki on yksi onnistuneen datanhallinnan edellytyksistä, sillä datanhallinnan kehitys vaatii usein organisaation laajuisia muutoksia, joiden läpiviemiseksi tarvitaan johdon tuki ja hyväksyntä (Mahanti 2018; Chakravorty 2020). On siis olennaista, että organisaation johto ymmärtää datanhallinnan tärkeyden.

Tutkimuksessa kävi ilmi myös kohdeorganisaation vahva pyrkimys liiketoimintalähtöisyyteen. Tämä pyrkimys kävi ilmi erityisesti data- ja analytiikkakehityksessä. Haastateltavat kokivat, että nykyisellä organisoitumistavalla kehitystä tehdään vahvasti liiketoiminnan tarpeisiin. Vaikka haastateltavat eivät nähneet yhtä vahvaa yhteyttä datanhallinnan ja liiketoiminnan tarpeiden välillä, datanhallinnan linjaus data- ja analytiikkakehitysvastuun hajauttamisesta liiketoimintayksiköille edustaa kohdeorganisaation pyrkimystä liiketoimintalähtöisiin toimintatapoihin myös datanhallinnassa. Empiirisessä tutkimuksessa pystyttiin myös tunnistamaan yhteys datanhallinnan ja liiketoiminnan tarpeiden välillä. Teoreettisessa viitekehyksessä tunnistettiin, että organisaation datanhallinta tulisi suunnitella tukemaan liiketoimintastrategiaa ja -tavoitteita (DAMA 2017; Abraham et al. 2019; Chakravorty 2020; Lis et al. 2023). Tämän vuoksi pyrkimys tehdä asioita liiketoimintalähtöisesti luokiteltiin kohdeorganisaation datanhallinnan vahvuudeksi. Kohdeorganisaation datanhallinnan ja liiketoiminnan yhteyttä voisi kuitenkin vielä vahvistaa kiinnittämällä huomiota datanhallinnan toteutuksen edellytyksiin sekä datanhallinnan ja organisaation strategian väliseen yhteyteen.

Datan ja datanhallinnan arvon ymmärtämisen sekä liiketoimintalähtöisyyden lisäksi kohdeorganisaation datanhallinnan vahvuudet liittyvät datanhallinnan aktiiviseen kehityk-

seen ja kontekstitekijöiden kanssa ristiriidattomaan datanhallintamalliin. Kohdeorganisaatio on kehittänyt viime vuosina datanhallintamalliaan luomalla tietoaaluejakoja ja kehittämällä datanhallinnan rooleja. Empiirisessä tutkimuksessa ei noussut esiin ristiriitoja näiden ja kontekstitekijöiden välillä. Yhtiössä arvostetaan joustavuutta päätöksenteossa, mutta toisaalta toimiala asettaa vaatimuksia säännöstenmukaisuuden ja turvallisuuden suhteen. Haastatteluissa ei käynyt ilmi, että datanhallinta ei tukisi näitä tarpeita. Vaikka teoreettisessa viitekehyksessä yhtiön kokoa ei vahvistettu varsinaiseksi kontekstitekijäksi, ristiriitoja ei tunnistettu myöskään organisaation koon ja datahallintamallin välillä. Tämä oli todennäköisesti sen ansiota, että kohdeorganisaatiossa oli jo aiemmin reagoitu data- ja henkilöstömäärän kasvuun ja niiden luomaan hallinnan tarpeeseen kehittämällä määritellympiä datanhallinnan rakenteita. Teoreettisessa viitekehyksessä todettiin, että kontekstitekijöiden huomioiminen mahdollistaa organisaatioille datanhallinnan muokkauksen omiin ympäristöihin ja tarpeisiin sopiviksi (Weber et al. 2009; Otto 2011b; Abraham et al. 2019; Lis et al. 2023). Tämän vuoksi ristiriidattomuus kontekstitekijöiden kanssa nähtiin vahvuutena. Ristiriidattomuudesta voidaan olettaa, että datanhallintamalli sopii kohdeorganisaation toimintaan. Tässä tulee kuitenkin huomioida, että empiirisessä tutkimuksessa tunnistettiin, ettei kohdeorganisaation datanhallintamalli ollut täysin jalkautunut. On mahdollista, että paremman jalkautuksen myötä ilmenee vielä ristiriitoja, jotka eivät olleet tutkimuksen toteutuksen aikaan havaittavissa.

Kohdeorganisaation datanhallinnan viimeinen tunnistettu vahvuus oli datan hyödyntämistä ja hallintaa tukevat teknologiat. Empiirisessä tutkimuksessa todettiin, että kohdeorganisaatiossa on panostettu datan hyödyntämistä ja hallintaa tukevien teknologioiden hankkimiseen. Tutkimuksessa kävi ilmi, että teknologiat eivät muodosta organisaatiossa hidastetta, vaan niiden puolesta liiketoiminnan tarpeet pystytään hyvin täyttämään.

Edellä esiteltyjen vahvuuksien perusteella voidaan todeta, että kohdeorganisaation datanhallintamalli tukee organisaation toimintaa melko hyvin, sillä sen taustalla vaikuttavat liiketoiminnan tarpeet ja pyrkimys tehdä datanhallintaa liiketoimintalähtöisesti. Datanhallintamallin ja organisaatiokohtaisten kontekstitekijöiden välillä ei myöskään todettu ristiriitaisuuksia, jotka vähentäisivät mallin sopivuutta organisaation toimintaan tai aiheuttaisivat liiketoiminnalle haasteita. Lisäksi mallia kehitetään aktiivisesti, joten organisaatiossa ei suhtauduta datanhallintaan vain kertaluontoisena projektina. Tämä mahdollistaa mallin muokkaamisen organisaation muuttuvien tarpeiden mukaan. Organisaation toimintaa tukevat myös erilaiset teknologiat, jotka mahdollistavat osaltaan datan tehokkaan hallinnan ja hyödyntämisen. Näistä huolimatta kohdeorganisaation datanhallinnassa voitiin tunnistaa myös kehityskohteita, joita edistämällä voitaisiin todennäköisesti lisätä datanhallinnan positiivisia vaikutuksia.

Kohdeorganisaation datanhallinnan keskeisimmiksi kehityskohteiksi tunnistettiin käytäntöjen ja rakenteiden jalkautus, viestintä, yhteistyö ja yhteiset toimintatavat, aikaresurssit ja datanhallinnan hyötyjen konkretisointi (taulukko 18). Kehityskohteet tunnistettiin taustoitavan kyselyn ja haastatteluiden avulla.

Taulukko 18. *Keskeisimmät tunnistetut kehityskohteet kohdeorganisaation datanhallinnassa*

Kehityskohde	Kuvaus
Käytäntöjen ja rakenteiden jalkautus	Kohdeorganisaatiossa ylätasolla suunnitellut asiat eivät aina jalkaudu käytännön toimintaan. Esimerkiksi datanhallinnan roolit ja niiden vastuut saattavat olla rooleihin nimetyille henkilöille epäselviä.
Viestintä ja osaaminen	Tietoisuus datanhallinnasta ja sen politiikoista ja rakenteista ei saavuta kaikkia. Vaikka kaikki henkilöt eivät tarvitse kaikkea tietoa datanhallinnasta, yleisen tietoisuuden ja osaamisen lisäämiselle nähdään tarve.
Yhteistyö ja yhteiset toimintatavat	Eri tiimit eivät tee tarpeeksi yhteistyötä, jolloin päädytään väliillä tekemään päällekkäisiä asioita. Hyviä käytänteitä jaetaan tiimien välillä harvoin ja yhteisiä toimintatapoja on sovittu vain vähän.
Aikaresurssit	Datanhallinnan rooleja omaavilla henkilöillä ei ole tarpeeksi aikaa hoitaa vastuitaan. Lisäksi liiketoimintalähtöinen data- ja analytiikkakehitys vaatii resursseja liiketoiminnalta. Tällä hetkellä liiketoiminnalla ei aina ole tarpeeksi aikaa osallistua data- ja analytiikkakehitykseen.
Datanhallinnan hyötyjen konkretisointi	Hyötyjen konkretisoinnissa nähtiin mahdollisuuksia lisätä henkilöstön sitoutumista datanhallintaan.

Empiirisessä tutkimuksessa tunnistettiin selkeä tarve jalkauttaa paremmin kohdeorganisaation datanhallinnan rakenteet ja periaatteet. Tämä tarve ilmeni erityisesti datanhallinnan roolien ja vastuiden yhteydessä. Tutkimuksessa todettiin, että kohdeorganisaatiossa on määriteltyjä datanhallinnan rooleja ja niistä vastaavia henkilöitä, mutta jopa näille henkilöille voi olla epäselvää, mitä vastuita heille kuuluu. Datanhallinnan roolien omistajilla ei myöskään aina ole ollut todellista mahdollisuutta toteuttaa vastuitaan, sillä tehtäville ei ole varattu tarpeeksi työaika. Datanhallinnan roolit ja vastuut olivat epäselviä myös muulle organisaatiolle, minkä takia asioiden vastuuhenkilöiden löytäminen ei aina ole ollut helppoa. Roolien ja vastuiden jalkauttamisen lisäksi haasteita todettiin datanhallinnan periaatteiden jalkauttamisessa. Datanhallinnan periaatteiden todettiin jäävät usein vain ylätasoin linjauksiksi, joista niiden toteuttamisesta vastaavat henkilöt eivät aina tiedäneet, miten niiden tulisi näkyä käytännön tekemisessä. Tutkimuksessa todettiin, että haasteet roolitusten ja periaatteiden jalkauttamisessa näkyivät eniten asiantuntijoille ja tiiminvetäjille. Tämä selittyi sillä, että epäselvien vastuiden vaikutuksia ilmeni lähinnä

käytännön toiminnan tasolla: asioita saattoi jäädä tekemättä tai tekemisen edistäminen oli hitaampaa, koska ei ollut selkeää, keneltä tukea voi pyytää.

Teoreettisessa viitekehysessä todettiin, että selkeät vastuut ja roolit ovat yksiä datanhallinnan onnistumisen kannalta kriittisiä tekijöitä (Mahanti 2018; Al-Ruithe et al. 2019; Alhassan et al. 2019). Erityisesti datanhallinnan roolien selkeyttäminen ja jalkauttaminen olisi siis kohdeorganisaatiossa tärkeää. Tätä voitaisiin tehdä esimerkiksi kartoittamalla tarvittavat roolit ja niihin liittyvät vastuut ja arvioimalla, kuinka paljon aikaa vastuiden toteuttamiseen tarvitaan. Tämän jälkeen tulisi käydä läpi, onko määriteltyihin rooleihin nimetty jo vastuuhenkilöt. Vastuuhenkilöiden kanssa tulisi käydä keskusteluja siitä, ovatko tehtävän vastuut selkeät, onko henkilöillä tarpeeksi osaamista vastuiden täyttämiseen ja onko heidän resursoinnissaan varattu tarpeeksi aikaa tehtävien toteuttamiseen. Jos näin ei ole, tulee resursointia tarkastella uudelleen tai määrätä rooleihin uusia vastuuhenkilöitä. Tällä tavalla varmistetaan, että datanhallinnan rooleihin ei ole vain kirjattu jonkun nimeä, vaan vastuuhenkilöt ovat tietoisia rooleistaan ja vastuistaan ja heillä on todellinen mahdollisuus täyttää velvoitteensa.

Empiirisessä tutkimuksessa tunnistettiin myös tarve kehittää datanhallintaa koskevaa viestintää ja osaamista. Tällä hetkellä tieto datanhallinnasta ja sen politiikoista ei saavuttanut kaikkia oleellisia henkilöitä, mutta toisaalta liiketoimintaa ei myöskään haluttu kuormittaa turhalla informaatiolla. Henkilöstöllä todettiin siis olevan erilaisia viestintä- ja osaamistarpeita, jotka tulisi ottaa huomioon. Viestinnän ja osaamisen kehitystarpeita tunnistettiin erityisesti datakehityksen ulkokehän sidosryhmiin liittyen. Heidän tietoisuuttaan tulisi lisätä ainakin datanhallinnan vastuurakenteista, joita pidettiin tällä hetkellä epäselvinä. Kohdeorganisaatiossa tulisi myös varmistaa, että kaikki sidosryhmät ovat tarvittavilta osin tietoisia heitä koskevista datanhallinnan politiikoista. Tämä on tärkeää, sillä jos sidosryhmät eivät ole tietoisia heitä koskevista odotuksista, ei toiminnassa tapahdu muutoksia. Myös teoreettisessa viitekehysessä todettiin, että tehokkaalla kommunikoinnilla ja viestinnällä on iso merkitys sen varmistamisessa, että sidosryhmät tietävät mitä heiltä odotetaan datanhallintaan liittyen (DAMA 2017; Mahanti 2018). Kohdeorganisaatiossa voitaisiin lähteä kehittämään viestintää määrittelemällä ensiksi, millaisia sidosryhmiä datanhallinnalla on ja mitä heidän viestinnälliset tarpeensa ovat. Tämän pohjalta voitaisiin lähteä laajentamaan henkilöstön tietämystä ja osaamista datanhallinnasta olennaisilta osin, kuormittamatta kuitenkaan liikaa liiketoimintaa. Teoreettisen viitekehityksen mukaan tässä voidaan käyttää apuna viestintäsuunnitelmaa, joka kattaa keinot viestiä miksi ja miten datanhallintaa tehdään ja mitä hyötyjä datanhallinta tuottaa (Mahanti 2018; Abraham et al. 2019).

Kolmas empiirisessä tutkimuksessa tunnistettu kehityskohde oli yhteistyö ja yhteiset toimintatavat. Tutkimuksessa todettiin, että yhteistyö liiketoiminnan ja data- ja analytiikkakehityksen välillä toimii pääosin hyvin, mutta yleisesti ottaen yhteistyötä eri tiimien välillä kaivattaisiin datanhallinnassa enemmän. Säännöllisempi yhteistyö tiimien välillä voisi parantaa tilannekuvaa isojen, eri tiimejä yhdistävien kehitysprojektien osalta ja edistää hyvien käytäntöjen leviämistä. Lisäksi yhteistyötä lisäämällä voitaisiin vähentää päällekkäisen työn tekemistä. Tutkimuksessa nimittäin todettiin, että tiimit edistävät välillä samoja asioita toisista tietämättään. Tarve yhteistyön lisäämiselle nousi esiin erityisesti asian tuntijatasolta. Myös työn teoreettinen viitekehys tukee yhteistyön tärkeyttä kehityskohdeena. Teoreettisessa viitekehyksessä todettiin, että tiimien välinen yhteistyö edistää omistajuutta, kommunikaatiota ja koko organisaation kattavan kuvan muodostumista (Mahanti 2018).

Yhteistyön niukkuus ilmeni myös yhteisten käytäntöjen pienenä määränä kohdeorganisaation data- ja analytiikkakehityksessä. Tutkimuksessa kävi ilmi, että kohdeorganisaatiossa on ollut tyypillistä keskittyä enemmän asioiden edistämiseen kuin niiden raamittamiseen. Halu edistää asioita sekä jo aiemmin todettu ajan puute ovat osaltaan varmasti vaikuttaneet siihen, että asioita mieluummin viedään eteenpäin, kuin pysähdytään sopimaan yhteisiä käytänteitä. Kuitenkin tutkimuksessa tunnistettiin, että määritellyistä toimintatavoista voisi olla hyötyä esimerkiksi datan laadun ja raporttien toimivuuden seurannassa sekä raporttitarpeiden hallinnassa. Tällä hetkellä vasta loppukäyttäjä saattoi huomata, jos data ei ollut päivittynyt tai raportti oli mennyt rikki. Uusia raportteja oli myös pyydetty matalalla kynnyksellä, mikä oli johtanut siihen, että kaikille raporteille ei ollut todellisuudessa käyttöä. Kohdeorganisaatiossa voitaisiin siis muun muassa vähentää loppukäyttäjälle näkyviä virheitä ja saavuttaa tehokkuutta määrittelemällä yhteisiä toimintatapoja tai toimintatapojen raameja.

Kohdeorganisaatio voisi lisätä tiimien välistä yhteistyötä lisäämällä mahdollisuuksia tiimien väliseen vuorovaikutukseen esimerkiksi tiimien yhteisten tilaisuuksien avulla. Tilaisuuksissa voitaisiin esimerkiksi jakaa ajankohtaista tietoa ja hyväksi todettuja käytänteitä. Tilaisuuksissa voitaisiin myös sopia yhteisistä menettelytavoista ja käydä läpi, miten eri tiimit toteuttavat datanhallinnan politiikoita, joiden vienti käytäntöön koettiin välillä hankalaksi.

Neljäs empiirisessä tutkimuksessa tunnistettu datanhallinnan kehityskohde oli aikaresurssit. Tutkimuksessa kävi ilmi, että kaikilla datanhallinnan rooleihin nimetyillä henkilöillä ei ole tarpeeksi aikaa toteuttaa vastuitaan, koska heidän pääasialliset työtehtävänsä vievät käytännössä kaiken työajan. Lisäksi data- ja analytiikkakehityksessä liike-

toiminnan resursoinnin ei todettu aina vastaavan liiketoimintalähtöisen data- ja analytiikkakehityksen tarpeita. Resursointia datanhallintaan voitaisiin parantaa jo aiemmin ehdotetuilla toimilla eli tehtäviin menevän työajan määrittelyllä ja keskusteluilla rooleista vastaavien henkilöiden kanssa. Lisäksi liiketoiminnan resursointia data- ja analytiikkakehitykseen tulisi tarkastella uudestaan. Olennaista olisi selvittää, minkälainen panostus liiketoiminnalta tarvitaan ja vastaako se heillä käytettävissä olevia resursseja. Jos näin ei ole, tulisi selvittää, miten liiketoiminnan resursointia data- ja analytiikkakehitykseen voitaisiin lisätä vai siedetäänkö joissain tapauksissa kehityksen heikompa yhteyttä liiketoimintaan.

Viimeinen empiirisessä tutkimuksessa tunnistettu kehityskohde oli datanhallinnan hyötyjen konkretisointi. Vaikka kohdeorganisaation yksi vahvuus oli, että kohdeorganisaatiossa ymmärretään datanhallinnan merkitys ja sen nähdään tuottavan arvoa, tutkimuksen mukaan kaikille ei ole täysin selkeää, mitä arvoa datanhallinta tuottaa. Tutkimuksessa kävi lisäksi ilmi, että yhtiössä keskitytään usein mieluummin asioiden edistämiseen kuin niiden raamittamiseen. Jos raamittaminen jää usein tekemisen jalkoihin, se voi olla käytännön osoitus siitä, että datanhallinnan hyödyt eivät ole kaikille täysin selkeitä. Haastateltavat pystyivät kuitenkin tunnistamaan joitain datanhallinnan hyötyjä. Esiin tuodut hyödyt liittyivät pääasiassa tehokkuuden lisäämiseen ja datan hyödynnettävyyden parantamiseen. Näkemyksissä datanhallinnan hyötyihin liittyen ei voitu erottaa selkeitä eroja eri organisaatiotasojen välillä.

Konkretisoimalla datanhallinnan hyötyjä kohdeorganisaatiossa voitaisiin muistuttaa datanhallinnan oleellisuudesta ja mahdollisesti nostaa sen prioriteettia kiireen keskellä. Teoreettisessa viitekehityksessä todettiin, että on tärkeää saada kaikki sidosryhmät vakuuttamaan datanhallinnan arvosta, sillä datanhallinnan pitäisi olla koko organisaation yhteinen asia (Weber et al. 2009; Al-Ruithe et al. 2019; Tsavatewa 2023). Tämä kuitenkin edellyttää sitä, että organisaatiossa tiedetään, mitä hyötyjä datanhallinta luo (DAMA 2017; Mahanti 2018). Datanhallinnan hyötyjen konkretisointi on siis tärkeää datanhallinnan onnistumisen kannalta.

Teoreettisessa viitekehityksessä todettiin, että datanhallinnan tuottamaa arvoa voidaan arvioida muun muassa mittaamalla riskien vähentymistä ja operaatioiden tehokkuuden parantumista (DAMA 2017). Teoreettisessa viitekehityksessä kuitenkin todettiin myös datanhallinnan tavoitteiden vaikutus datanhallinnan arviointiin. Jos halutaan todistaa datanhallinnan onnistuminen, tulisi vaikutuksia peilata datanhallinnan tavoitteisiin ja johtaa seurattavat mittarit niistä. Kohdeorganisaation datanhallinnan visiona oli, että data löytyy luotettavasti, datan hyödyntäminen on helppoa ja dataratkaisuiden tekeminen on nopeaa. Taustoittavalla kyselyllä selvitettiin näiden tavoitteiden tilaa, jolloin todettiin, että

eniten kehitettävää oli datan löydettävyydessä. Kohdeorganisaatiossa voitaisiin seurata datanhallinnan onnistumista vastaavilla, mutta laajemmilla henkilöstölle lähetettävillä kyselyillä. Kyselyillä voitaisiin tiedustella organisaatiolle tärkeiden tekijöiden tilaa säännöllisin väliajoin. Tällöin voitaisiin konkretisoida datanhallinnan kehittämisen vaikutus ja seurata, miten erilaiset kehitystoimenpiteet vaikuttavat mittareihin. Lisäksi vaikutuksia voitaisiin konkretisoida esimerkkiprojektien kautta. Osoittamalla, miten projekti hyötyi datanhallinnasta, voidaan konkretisoida datanhallinnan vaikutuksia niiden asioiden kautta, jotka ovat liiketoiminnalle merkittäviä.

Tutkimuksessa tunnistetuista kohdeorganisaation vahvuuksista ja kehityskohteista nähdään, että vahvuudet liittyvät pitkälti kuvan 17 kehitysmallin alkupäähän eli datanhallintamallin luomiseen ja kehityskohteet datanhallinnan toteuttamiseen ja arviointiin. Hyödyntämällä tutkimuksessa luotua kehitysmallia kohdeorganisaatio voi varmistaa, että datanhallinnan kehitys ei jää vain datanhallintamallin uudistamiseen, vaan myös datanhallinnan toteutus ja arviointi tulee olennaiseksi osa datanhallinnan kehitystä. Tämä datanhallinnan toteutuksen ja arvioinnin vahvempi huomioiminen auttaisi kohdeorganisaatiota edistämään tutkimuksessa tunnistettuja datanhallinnan kehityskohteita.

Lisäksi kehitysmallin hyödyntäminen varmistaa, että kohdeorganisaation vahvuudet pysyvät vahvuuksina jatkossakin, vaikka organisaation tavoitteissa tai toimintoympäristössä tapahtuisi muutoksia. Kiinnittämällä huomioita organisaation datanhallinnan vahvuuksiin, niitä voidaan myös vahvistaa edelleen. Yksi tutkimuksessa tunnistettu vahvuus oli esimerkiksi datanhallintamallin ristiriidattomuus kontekstitekijöiden kanssa. Panostamalla mallin määrittelyyn kontekstitekijöiden mukaiseksi, voitaisiin mallin yhteensopiavuutta kontekstitekijöiden kanssa vielä vahvistaa. Näillä tavoin kohdeorganisaatio voisi mallin avulla uudistaa datanhallintaansa tukemaan vielä paremmin organisaation toimintaa.

6. YHTEENVETO

Tässä luvussa tehdään tutkimuksesta yhteenveto ja vastataan tutkimuskysymyksiin. Lopuksi vielä arvioidaan tutkimusta ja esitetään jatkotutkimusmahdollisuuksia.

6.1 Tutkimuskysymyksiin vastaaminen

Tämän tutkimuksen päätutkimuskysymys oli ”Miten organisaatio voi kehittää datanhallintaa omaan toimintaansa sopivaksi?”. Kysymykseen pyrittiin vastaamaan alatutkimuskysymysten avulla, joista ensimmäinen oli ”Miten datanhallintamalli voidaan muodostaa tai uudistaa tietyille organisaatiolle sopivaksi?”. Tutkimuksessa tunnistettiin, että organisaatiokohtaisen ja organisaation toimintaa tukevan datanhallintamallin muodostamisessa tai uudistamisessa olennaista on organisaation liiketoimintastrategian jalostaminen datanhallinnan tavoitteiksi ja näiden tavoitteiden huomioiminen hallintamallin muodostamisessa. Lisäksi tutkimuksessa tunnistettiin muiden kontekstitekijöiden huomioimisen tärkeys. Näitä olivat lait ja sääntely, toimiala, toimintaympäristön muutosnopeus, organisaation rakenne, kilpailustrategia, prosessien harmonisointiaste, päätöksentekotyylit ja kulttuuri, datan tyyppi ja kasvuvauhti ja organisaation toiminnan laajuus. Datanhallintamalli tulisi muodostaa siten, että se on yhteensopiva näiden organisaatiokohtaisten tekijöiden kanssa. Tutkimuksessa tunnistettiin myös edellä esiteltyjen tekijöiden ja datanhallintamallin yhteensopivuuden jatkuvan arvioinnin oleellisuus. Datanhallinta tunnistettiin jatkuvaksi prosessiksi, jota tulisi uudistaa sisäisten ja ulkoisten muutosten mukaan.

Tutkimuksen toinen alatutkimuskysymys oli ”Mitkä asiat tukevat datanhallinnan toteuttamista?”. Tutkimuksessa tunnistettiin useita datanhallinnan edellytyksiä, joiden huomiotta jättäminen voi johtaa datanhallinnan epäonnistumiseen. Näitä olivat johdon tuki, rakenteiden ja periaatteiden jalkauttaminen, työntekijöiden sitouttaminen, henkilöstön kyvykkydet, viestintä ja yhteistyö ja aika- ja teknologioresurssit. Lisäksi tutkimuksessa tunnistettiin datanhallinnan arvioinnin merkitys. Datanhallinnan arvioinnilla voidaan konkretisoida datanhallinnan hyötyjä ja näin lisätä työntekijöiden sitoutumista datanhallintaan. Lisäksi datanhallinnan arvioinnilla voidaan seurata, saadaanko datanhallinnalla aikaan tavoiteltuja vaikutuksia.

Työn viimeinen alatutkimuskysymys oli ”Mitkä ovat kohdeorganisaation nykyisen datanhallinnan vahvuudet ja heikkoudet?”. Kohdeorganisaation datanhallinnan keskeisimmiksi vahvuuksiksi tunnistettiin ymmärrys datan arvosta liiketoiminnalle, ymmärrys da-

tanhallinnan arvosta liiketoiminnalle, johdon tuki, pyrkimys liiketoimintalähtöisyyteen, datanhallinnan rakenteiden aktiivinen kehitys, kontekstitekijöiden kanssa ristiriidaton datanhallintamalli sekä datan hyödyntämistä ja hallintaa tukevat teknologiat. Keskeisimmiksi kehityskohteiksi taas tunnistettiin käytäntöjen ja rakenteiden jalkautus, viestintä, yhteistyö ja yhteiset toimintatavat, aikaresurssit ja datanhallinnan hyötyjen konkretisointi.

Päätutkimuskysymykseen saatiin vastattua muodostamalla näiden alatutkimuskysymysten vastausten avulla malli datanhallinnan kehittämiseen organisaation toimintaan sopivaksi. Mallin rakenne pohjautuu PDCA-malliin, joten se tukee jatkuvaa parantamista ja sisäisiin ja ulkoihin muutoksiin reagoimista. Mallin vaiheiksi määriteltiin organisaatiokohtaisen datanhallintamallin muodostaminen tai uudistaminen, datanhallinnan toteuttaminen, datanhallinnan arviointi ja toiminnan vakiinnuttaminen tai muutosten tekeminen. Malli on esitelty kuvassa 17. Mallia peilattiin vielä kohdeorganisaation keskeisimpiin datanhallinnan vahvuuksiin ja heikkouksiin. Analysoimalla kohdeorganisaation keskeisimpiä datanhallinnan vahvuuksia ja heikkouksia todettiin, että ottamalla käyttöön tutkimuksessa kehitetyn datanhallinnan kehitysmallin, kohdeorganisaatio voisi keskittää huomionsa datanhallinnan toimivuuden kannalta olennaisiin asioihin, vahvistaa edelleen datanhallinnan vahvuuksiaan sekä parantaa tunnistettuja kehityskohteita. Näin kohdeorganisaatio voisi uudistaa datanhallintaansa tukemaan vielä paremmin organisaation toimintaa.

6.2 Tutkimuksen arviointi

Laadullista tutkimusta voidaan arvioida Shentonin (2004) määrittelemän neljän ulottuvuuden avulla. Nämä ulottuvuudet ovat uskottavuus, siirrettävyys, luotettavuus ja vahvistettavuus. Tutkimuksen uskottavuudella arvioidaan, kuinka yhteneväisiä löydökset ovat todellisuuden kanssa. Tutkimuksen uskottavuutta voidaan lisätä muun muassa löydösten peilaamisella aikaisempaan tutkimukseen, tutkijan perehtymisellä kohdeorganisaation kulttuuriin ja luottamuksen rakentamisella haastateltavien kanssa. (Shenton 2004) Tässä työssä empiirinen tutkimus tehtiin yksittäisen tapauksen tapaustutkimuksena, mutta sen uskottavuutta lisättiin peilaamalla tuloksia teoreettiseen viitekehykseen. Teoreettinen viitekehys rakennettiin kirjallisuuskatsauksen avulla, jonka uskottavuutta pyrittiin lisäämään kuvaamalla aineiston valinta tarkasti ja laittamalla lähdeviittaukset selkeästi näkyviin.

Tutkimuksen uskottavuutta lisäsi tutkijan työskentely kohdeorganisaatiossa, koska sen ansiosta organisaation kulttuuri oli tutkijalle jo entuudestaan tuttu. Toisaalta tutkijan työskentely kohdeorganisaatiossa saattoi heikentää tutkijan objektiivisuutta, sillä hänelle oli

muodostunut jo jonkinlainen käsitys organisaation datanhallinnasta. Työsuhteen ansiosta osa haastateltavista oli tutkijalle jo tuttuja, joten he pystyivät puhumaan tutkimuskohteesta hyvin avoimesti. Muiden kanssa luottamusta ja rehellisyyttä pyrittiin lisäämään kysymällä halukkuutta osallistua haastatteluun, ilmoittamalla tulosten anonyymisoimisesta ja kiinnittämällä huomiota haastattelun ilmapiiriin. Tutkimuksen uskottavuutta voi kuitenkin heikentää haastateltavien harkinnanvarainen otanta. Shentonin (2004) mukaan sattumanvarainen otanta estää tutkijan ennakko-oletuksien vaikutuksen otantaan. Tutkimuksessa päädyttiin harkinnanvaraiseen otantaan, sillä sen avulla varmistettiin, että haastateltavilla oli riittävästi näkemyksiä ja kokemuksia datanhallinnasta.

Tutkimuksen siirrettävyydellä arvioidaan laajuutta, jolla tutkimuksen tuloksia voidaan soveltaa muissa konteksteissa. Tutkimuksen siirrettävyyttä voidaan parantaa kuvaamalla selkeästi tutkimusasetelman taustatiedot, tutkimukseen vaikuttavat rajoitteet sekä tutkimusprosessi. (Shenton 2004) Tässä tutkimuksessa siirrettävyyttä pyrittiin parantamaan kuvaamalla työn tutkimusmetodologia Saundersin ja muiden (2019) tutkimussipulin avulla. Tutkimuksen elementit eli tutkimussipulin eri kerrokset kuvattiin tarkasti. Erityistä huomiota kiinnitettiin aineiston keruun käytännön tason kuvaamiseen. Kirjallisuuskatsauksen kuvauksessa tuotiin ilmi käytetyt hakusanat ja tietokannat sekä aineiston valinnan perustelut. Haastatteluiden osalta työssä tuotiin ilmi muun muassa haastattelukysymykset, haastateltavien määrä, roolit, haastatteluiden kesto ja aikaväli, jolla haastattelut toteutettiin. Lisäksi kohdeorganisaation esittelyssä tuotiin esiin oleelliset taustatiedot kohdeorganisaatiosta ja sen datanhallinnan tilasta. Tutkimuksen tuloksia sovellettaessa täytyy kuitenkin huomioida, että tutkimus toteutettiin yksittäisen tapauksen tapaustutkimuksena, joten sen tuloksia ei voida välttämättä yleistää kaikissa ympäristöissä.

Tutkimuksen luotettavuus kuvaa tutkimuksen johdonmukaisuutta ja kuinka hyvin tutkimus on toistettavissa tutkimusprosessin kuvauksen perusteella. Tutkimuksen luotettavuutta voidaan lisätä kuvaamalla tutkimusprosessi ja datan keruu huolellisesti sekä arvioimalla tutkimusprosessin tehokkuutta. (Shenton 2004) Kuten jo aiemmin todettu, tässä työssä pyrittiin kuvaamaan tutkimusprosessi huolellisesti ja aineiston keruu kuvattiin käytännön tasolla. Aineistonkeruumenetelmänä oli puolistrukturoidut haastattelut, joten haastateltavat saivat vastata kysymyksiin vapaasti. Tästä syystä haastateltaville saatettiin esittää erilaisia tarkettavia kysymyksiä, joita ei ole dokumentoitu. Tämä saattaa vaikuttaa siihen, että toteutettaessa vastaava tutkimus samoilla menetelmillä ja samassa ympäristössä ei välttämättä saataisi täysin vastaavia tuloksia.

Tutkimuksen vahvistettavuus kuvaa tutkijan objektiivisuutta ja sitä, miten hyvin työn tulokset perustuvat tutkimuskohteiden näkemyksiin. (Shenton 2004) Tämän tutkimuksen

objektiivisuutta pyrittiin lisäämään haastatteluiden aikana kysymällä tarkentavia kysymyksiä ja toistamalla ääneen haastateltavien näkemyksiä, jotta voitiin varmistaa, että haastateltavan näkökulma oli ymmärretty oikein. Haastattelut myös tallennettiin ja litte- roitiin, jolloin tutkija pystyi perustamaan analyysinsä koko haastatteluaineistoon, eikä vain omiin muistiinpanoihinsa ja muistiinsa luottaen. Tämä mahdollisti myös suorien lainausten poimimisen. Suorien lainauksien esittelyllä voitiin lisätä vastausten analysoinnin luotettavuutta. Vaikka tutkimuksessa pyrittiin objektiivisuuteen, tutkijan subjektiiviset näkemykset ovat vaikuttaneet esimerkiksi siihen, mitä tutkimuksia kirjallisuuskatsaukseen valittiin. Tehdyt valinnat pyrittiin kuitenkin perustelemaan läpinäkyvyyden lisäämiseksi. Lisäksi tutkimuksen objektiivisuutta saattaa heikentää jo aiemmin ilmi tuotu tutkijan työskentely kohdeorganisaatiossa. Työskentelyn kautta kerrytetty kokemus ja muodostetut näkemykset ovat saattaneet vaikuttaa tutkimuksen toteuttamiseen.

Tutkimuksen tavoitteena oli tuottaa tietoa datanhallinnan viitekehyksistä, sekä tekijöistä, joihin organisaatioiden tulisi kiinnittää huomiota muodostaessaan tai kehittäessään toimintaansa sopivaa datanhallintaa. Lisäksi tutkimuksessa haluttiin tuottaa tietoa kohdeorganisaation datanhallinnan nykytilasta sekä kehityskohteista. Tutkimuksen tavoitteet saavutettiin hyvin. Työssä pystyttiin luomaan jatkuvan kehityksen malli, jonka avulla organisaatiot voivat kehittää datanhallintaa toimintaansa sopivaksi. Malli kattaa kontekstitekijät, jotka huomioimalla organisaatiot voivat muodostaa omaan ympäristöönsä sopivan datanhallintamallin. Lisäksi malli huomioi datanhallinnan toteuttamisen ja sen edellytykset. Aikarajoitusten vuoksi mallia ei kuitenkaan keretty testaamaan muissa ympäristöissä, joten varmuudella ei voida sanoa, että sitä voitaisiin soveltaa muissa kuin kohdeorganisaation kaltaisissa yrityksissä. Datanhallinnan kehitysmallin luomisen lisäksi tutkimuksessa onnistuttiin kartoittamaan kohdeorganisaation datanhallinnan vahvuudet ja keskeisimmät kehityskohteet ja niitä voitiin peilata luotuun kehitysmalliin. Näin selvisi, että luodulla kehitysmallilla voitaisiin kehittää kohdeorganisaation datanhallintaa vielä paremmin organisaation toimintaan sopivaksi. Tutkimuksen tavoitteiden täyttämisen lisäksi tutkimuksella pystyttiin tuottamaan tietoa datanhallinnan käytännön soveltamisesta ja toteuttamisesta eli sillä onnistuttiin täydentämään johdannossa tunnistettua tutkimusaukkoa datanhallinnan käytännön toteuttamiseen liittyen.

6.3 Jatkotutkimusmahdollisuudet

Tutkimuksessa tunnistetut jatkotutkimusmahdollisuudet liittyvät muodostettuun datanhallinnan kehitysmalliin. Aikarajoitusten vuoksi mallin toimivuutta ei keretty testaamaan muissa ympäristöissä. Jatkotutkimuksessa voitaisiin testata, sopiiko malli hyödynnettä-

väksi muunlaisissa ympäristöissä ja ovatko empiirisen tutkimuksen avulla lisätyt kontekstitekijät ja datanhallinnan toteuttamisen edellytykset relevantteja myös kohdeorganisaation ulkopuolella. Lisäksi jatkotutkimuksessa voitaisiin selvittää, ovatko kaikki mallin osat yhtä olennaisia vai voiko niitä priorisoida joillakin perusteilla. Olisi esimerkiksi hyödyllistä selvittää, onko kontekstitekijöistä tai datanhallinnan toteuttamisen edellytyksistä jokin painoarvoltaan merkittävämpi kuin muut. Tällöin organisaatiot voisivat priorisoida datanhallinnan kehityksessään mallin olennaisimpia tekijöitä.

Tässä työssä datanhallinnan vaikutusten arviointiin ei onnistuttu kehittämään erityisen konkreettisia mittareita. Datanhallintaa toteuttavat yritykset voisivat kuitenkin hyötyä tutkimuksesta, jossa niitä selvitettäisiin tarkemmin. Jatkotutkimuksessa voitaisiin selvittää, miten datanhallinnan vaikutuksia kannattaa arvioida ja miten vaikutuksia voitaisiin konkretisoida käytännön tasolla.

LÄHTEET

Abraham, R., Schneider, J., & Vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, vol.49, pp. 424–438. <https://doi.org/10.1016/j.ijinfomgt.2019.07.008>

Abraham, R., Schneider, J., & vom Brocke, J. (2023). A taxonomy of data governance decision domains in data marketplaces. *Electronic Markets*, vol, 33(1). <https://doi.org/10.1007/s12525-023-00631-w>.

Abueed, R. A. I., & Aga, M. (2019). Sustainable knowledge creation and corporate outcomes: Does Corporate Data Governance Matter? *Sustainability*, vol.11(20). <https://doi.org/10.3390/su11205575>.

Ahmadi, S., Tavana, M. M., Shokouhyar, S., & Dortaj, M. (2022). A new fuzzy approach for managing data governance implementation relevant activities. *TQM Journal*, vol.34(5), pp. 979–1012. <https://doi.org/10.1108/TQM-01-2021-0015>.

Ai, B., Li, Y., Ma, W., Qiu, M. & Liu, M. (2023). Assessing the Value of Data Assets: An Exploratory Study of Valuation Methods. *IEEE 22nd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, pp. 2147-2151. doi: 10.1109/TrustCom60117.2023.00299.

Alhassan, I., Sammon, D. and Daly, M. (2018). Data governance activities: a comparison between scientific and practice-oriented literature. *Journal of Enterprise Information Management*, Vol. 31(2), pp. 300-316. <https://doi-org.libproxy.tuni.fi/10.1108/JEIM-01-2017-0007>.

Alhassan, I., Sammon, D., & Daly, M. (2019). Critical success factors for data governance: a telecommunications case study. *Journal of Decision Systems*, vol. 28(1), pp. 41–61. <https://doi.org/10.1080/12460125.2019.1633226>.

Al-Ruithe, M. & Benkhelifa, E. (2017). Analysis and Classification of Barriers and Critical Success Factors for Implementing a Cloud Data Governance Strategy. *8th International Conference on Emerging Ubiquitous Systems and Pervasive Networks. Procedia Computer Science*, vol.113, pp.223-232. doi: 10.1016/j.procs.2017.08.352.

Al-Ruithe, M., Benkhelifa, E. & Hameed, K. (2019). A systematic literature review of data governance and cloud data governance. *Personal and Ubiquitous Computing*, vol. 23, pp. 839–859. <https://doi-org.libproxy.tuni.fi/10.1007/s00779-017-1104-3>.

- Altamimi, H., Liu, Q., & Jimenez, B. (2023). Not Too Much, Not Too Little: Centralization, Decentralization, and Organizational Change. *Journal of Public Administration Research and Theory*, vol. 33(1), pp. 170–185. <https://doi.org/10.1093/jopart/muac016>.
- Batini, C., Cappiello, C., Francalanci, C., & Maurino, A. (2009). Methodologies for data quality assessment and improvement. *ACM Computing Surveys*, vol. 41(3), pp. 1–52. <https://doi.org/10.1145/1541880.1541883>.
- Baxter, P. & Jack, S. (2008). Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers. *Qualitative Report*, vol. 13(4), pp. 544–559.
- Begg, C., & Caira, T. (2012). Exploring the SME Quandary: Data Governance in Practise in the Small to Medium-Sized Enterprise Sector. *Electronic Journal of Information Systems Evaluation*, vol. 15(1), pp. 1–12.
- Benfeldt, O. (2017). A comprehensive review of data governance literature. *Selected Papers of the IRIS*, vol.8. <http://www.nova.edu/ssss/QR/QR13-4/baxter.pdf>.
- Benfeldt, O., Persson, J.S. & Madsen, S. (2020). Data Governance as a Collective Action Problem. *Information System Frontiers*, vol.22, pp. 299–313. <https://doi-org.lib-proxy.tuni.fi/10.1007/s10796-019-09923-z>.
- Bento, P., Neto, M. & Côte-Real, N. (2022). How data governance frameworks can leverage data-driven decision making: A sustainable approach for data governance in organizations. *17th Iberian Conference on Information Systems and Technologies (CISTI)*, Madrid, Spain, pp. 1-5, doi: 10.23919/CISTI54924.2022.9866895.
- Capgemini Research Institute. (2020). The data-powered enterprise: Why organizations must strengthen their data mastery. <https://www.capgemini.com/research-and-insight/the-data-powered-enterprise/>. [Viitattu: 21.10.2024]
- Chakravorty, R. (2020). Common challenges of data governance. *Journal of Securities Operations & Custody*, vol.13(1), pp. 23–43. <https://doi.org/10.69554/GAKU4007>.
- Dallemule, L., & Davenport, T. (2017). What's Your Data Strategy? *Harvard business review*, vol. 95(3), pp. 112–121. Harvard Business School Publishing Corporation.
- DAMA International (2017). *Data Management Body of Knowledge. Second Edition*. Technics Publications, New Jersey.
- DeStefano, R. J., Tao, L., & Gai, K. (2016). Improving Data Governance in Large Organizations through Ontology and Linked Data. *IEEE 3rd International Conference on Cyber Security and Cloud Computing (CSCloud)*, pp. 279–284. <https://doi.org/10.1109/CSCloud.2016.47>.

- Devos, J., Van Landeghem, H., & Deschoolmeester, D. (2012). Rethinking IT governance for SMEs. *Industrial Management + Data Systems*, vol. 112(2), pp. 206–223. <https://doi.org/10.1108/02635571211204263>.
- Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *The Academy of Management Review*, vol. 14(4), pp. 532–550. <https://doi.org/10.2307/258557>.
- Eriksson, P., & Koistinen, K. (2005). Monenlainen tapaustutkimus. Kuluttajatutkimuskeskus.
- Eskola, J., & Suoranta, J. (1998). Johdatus laadulliseen tutkimukseen. Vastapaino.
- Fattah, I. A. (2024). The mediating effect of data literacy competence in the relationship between data governance and data-driven culture. *Industrial Management + Data Systems*, vol. 124(5), pp. 1823–1845. <https://doi.org/10.1108/IMDS-11-2023-0812>.
- Farquhar, J. D. (2012). Case study research for business. SAGE. <https://doi.org/10.4135/9781446287910>.
- Fingrid. Esittely. <https://www.fingrid.fi/sivut/yhtio/esittely/>. [Viitattu 15.3.2024]
- Fingrid-lehti. 2019. Energia-ala muuttuu megatrendien mukana. *Fingrid-lehti*, vol. 22. 1/2019. https://www.fingrid.fi/globalassets/dokumentit/fi/julkaisut/asiakaslehdet/fingrid-lehti_1_2019.pdf. [Viitattu 8.2.2025]
- Fingrid-lehti. 2024. Vihreä siirtymä tarvitsee verkkoja. *Fingrid-lehti*, vol. 27. https://www.fingrid.fi/globalassets/dokumentit/fi/julkaisut/asiakaslehdet/fingrid_2_2024_300524.pdf. [Viitattu 8.2.2025]
- Fink, A. (2005). Conducting research literature reviews : from the Internet to paper (2. ed.). Sage.
- Gaol, F., Siswanto, R., & Matsuo, T. (2023). Architectural modeling of data warehouse and analytic business intelligence for Bedstead manufacturers. *Open Engineering*, vol. 13(1), pp. 340–349. <https://doi.org/10.1515/eng-2022-0508>.
- Gupta, U. G., & Cannon, S. (2020). A Review of Data Governance Definitions and Emerging Perspectives. *International Journal of Data Analytics*, vol.1(2), pp. 30–47. <https://doi.org/10.4018/IJDA.2020070103>.
- Herselman, M., Wayi, N., & Olaitan, O. (2019). A Data Governance Maturity Evaluation Model for government departments of the Eastern Cape province, South Africa. *South African Journal of Information Management*, vol. 21(1), pp. 1–12. <https://doi.org/10.4102/sajim.v21i1.996>.

- Hirvonen, T. (2016). Tutkimusdatan hallinnan kehittäminen ja datan laadun arviointi asi-
antuntijayrityksessä. Diplomityö, Tampereen Yliopisto. <https://urn.fi/URN:NBN:fi:ttt-201606224303>.
- Hopper, M. A. (2023). Practitioner's guide to operationalizing data governance. Wiley-
Blackwell. <https://doi.org/10.1002/9781119851462>.
- Hyvärinen, M. K., Nikander, P., Ruusuvuori, J., & Aho, A. L. (2017). Tutkimushaastatte-
lun käsikirja. Vastapaino.
- Hyvärinen, M., Suoninen, E. & Vuori, J. (2021). Haastattelut. Teoksessa Jaana Vuori
(toim.) Laadullisen tutkimuksen verkkokäsikirja. Tampere: Yhteiskuntatieteellinen tieto-
arkisto. [https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/laadullisen-tutkimuk-
sen-aineistot/haastattelut/](https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/laadullisen-tutkimuk-sen-aineistot/haastattelut/). [Viitattu 8.2.2025]
- Isniah, S., Hardi Purba, H., & Debora, F. (2020). Plan do check action (PDCA) method:
literature review and research issues. Jurnal Sistem Dan Manajemen Industri, vol. 4(1),
pp. 72–81. <https://doi.org/10.30656/jsmi.v4i1.2186>.
- Jimenez, L., Polo, J., & Duarte, N. (2019). Overview of Data Governance in Business
Contexts. IOP Conference Series: Materials Science and Engineering, vol. 519(1).
<https://doi.org/10.1088/1757-899X/519/1/012023>.
- Jokela, E. (2023). Data governance -kypsytyksen arviointi ja -kehittäminen. Diplomityö,
Tampereen Yliopisto. <https://urn.fi/URN:NBN:fi:tuni-202306056502>.
- Juhila, K. (2021a). Koodaaminen. Teoksessa Jaana Vuori (toim.) Laadullisen tutkimuk-
sen verkkokäsikirja. Tampere: Yhteiskuntatieteellinen tietoarkisto. [Koodaaminen - Tieto-
arkisto](#). [Viitattu: 25.11.2024]
- Juhila, K. (2021b). Teemoittelu. Teoksessa Jaana Vuori (toim.) Laadullisen tutkimuksen
verkkokäsikirja. Tampere: Yhteiskuntatieteellinen tietoarkisto.
[https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/analyysitavan-valinta-ja-yleiset-
analyysitavat/teemoittelu/](https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/analyysitavan-valinta-ja-yleiset-analyysitavat/teemoittelu/). [Viitattu: 25.11.2024]
- Kallio, H., Pietilä, A.M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodo-
logical review: developing a framework for a qualitative semi-structured interview guide.
Journal of Advanced Nursing, vol. 72(12), pp. 2954–2965.
<https://doi.org/10.1111/jan.13031>.
- Karkošková, S. (2023). Data Governance Model To Enhance Data Quality In Financial
Institutions. Information Systems Management, vol. 40(1), pp. 90-110.
<https://doi.org/10.1080/10580530.2022.2042628>.

- Khatri, V. & Brown, C. (2010). Designing data governance. *Communications of the ACM*, vol. 53(1), pp. 148-152. doi: 10.1145/1629175.1629210.
- Kim, H. Y., & Cho, J. (2018). Data governance framework for big data implementation with NPS Case Analysis in Korea. *Journal of Business & Retail Management Research*, vol. 12(3), pp. 36–46. <https://doi.org/10.24052/jbrmr/v12is03/art-04>.
- Korhonen, J. J., Melleri, I., Hiekkänen, K., & Helenius, M. (2013). Designing Data Governance Structure: An Organizational Perspective. *GSTF Journal on Computing (JoC)*, vol. 2(4), pp. 11–17.
- Lee, Y., Madnick, S., Wang, R., Wang, F., & Zhang, H. (2014). A cubic framework for the chief data officer: Succeeding in a world of big data. *MIS Quarterly Executive*, vol. 13(1), pp. 1–13.
- Lis, D., Arbter, M., Spindler, M., & Otto, B. (2023). An Investigation of Antecedents for Data Governance Adoption in the Rail Industry-Findings From a Case Study at Thalys. *IEEE Transactions on Engineering Management*, vol. 70(7), pp. 2528–2545. <https://doi.org/10.1109/TEM.2022.3166109>.
- Mahanti, R. (2018). Data Governance Implementation: Critical Success Factors. *Software Quality Professional*, vol. 20(4), pp. 4–21.
- Mahanti, R. (2019). *Data Quality: Dimensions, Measurement, Strategy, Management, and Governance* (1st ed.). ASQ Quality Press.
- Martijn, N., Hulstijn, J., de Bruijne, M., & Tan, Y. (2015). Determining the Effects of Data Governance on the Performance and Compliance of Enterprises in the Logistics and Retail Sector. *Open and Big Data Management and Innovation. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol.9373, pp. 454–466. https://doi.org/10.1007/978-3-319-25013-7_37.
- Maulina, J., & Ruldeviyani, Y. (2019). Data Governance and Data Architecture for the Ministry of Foreign Affairs of the Republic of Indonesia. *2019 International Conference on Information Management and Technology (ICIMTech)*, pp. 100–414. <https://doi.org/10.1109/ICIMTech.2019.8843766>.
- Medeiros, M. M. de, Maçada, A. C. G., & Freitas Junior, J. C. da S. (2020). The effect of data strategy on competitive advantage. *The Bottom Line (New York, N.Y.)*, vol.33(2), pp. 201–216. <https://doi.org/10.1108/BL-12-2019-0131>.

- Medeiros, M., MaçAda, A., & Hoppen, N. (2021). The Role of Big Data Stewardship and Analytics as Enablers of Corporate Performance management. *RAM. Revista de Administração Mackenzie*, vol. 22(6), pp. 1–31. <https://doi.org/10.1590/1678-6971/eRAMD210063>
- Otto, B. (2011a). Data Governance. *Business & Information Systems Engineering*, vol.3(4), pp. 241–244. <https://doi.org/10.1007/s12599-011-0162-8>.
- Otto, B. (2011b). Organizing Data Governance: Findings from the telecommunications industry and consequences for large service providers. *Communications of the Association for Information Systems*, vol. 29(1), pp. 45–66. <https://doi.org/10.17705/1cais.02903>.
- Peng, G., Privette, J., Tilmes, C., Bristol, S., Maycock, T., Bates, J., Hausman, S., Brown, O., & Kearns, E. (2018). A Conceptual Enterprise Framework for Managing Scientific Data Stewardship. *Data Science Journal*, vol.17, pp. 15–15. <https://doi.org/10.5334/dsj-2018-015>.
- Romero, A., Gonzales, A., & Raymundo, C. (2019). Data Governance Reference Model under the Lean Methodology for the Implementation of Successful Initiatives in the Peruvian Microfinance Sector. *ACM International Conference Proceeding Series*, pp. 227–231. <https://doi.org/10.1145/3328833.3328859>.
- Saunders, M., Lewis, P., & Thornhill, A. (2019). *Research Methods for Business Students (Eighth Edition)*. Pearson Education, Limited.
- Schreibman, S., Siemens, R. & Unsworth, J. (2016). *A new companion to digital humanities*. 2nd ed. Wiley Blackwell.
- Shenton, A. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information*, vol. 22(2), pp. 63–75. <https://doi.org/10.3233/EFI-2004-22201>.
- Sothilingam, R., Pant, V., Shahrin, N., & Yu, E. (2021). Towards a Goal-Oriented Modeling Approach for Data Governance. *CEUR Workshop Proceedings*, vol.3045, pp. 69–77.
- Sulanen, S. (2021). Improving financial data quality through data governance. *Tampereen Yliopisto*. <https://urn.fi/URN:NBN:fi:tuni-202104263618>.
- Tan, K., Zhan, Y., Ji, G., Ye, F., & Chang, C. (2015). Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph. *International Journal of Production Economics*, vol.165, pp. 223–233. <https://doi.org/10.1016/j.ijpe.2014.12.034>.

- Tiilikainen, T. (2023). Enabling Data Platform Value Creation with Data Governance – A Case Study in Machinery Industry. Diplomityö, Tampereen Yliopisto. <https://urn.fi/URN:NBN:fi:tuni-202309218364>.
- Tsavatewa, C. (2023). Perceptions of data governance: identifying critical success factors in a university system's implementation effort. *Issues in Information Systems*, vol 24(4), pp. 390-411. http://doi.org/10.48009/4_iis_2023_130.
- Tuomi, J., & Sarajärvi, A. (2018). Laadullinen tutkimus ja sisällönanalyysi (Uudistettu laitotos.). Tammi.
- Voss, C., Tsikriktsis, N., & Frohlich, M. (2002). Case research in operations management. *International Journal of Operations & Production Management*, vol. 22(2), pp. 195–219. <https://doi.org/10.1108/01443570210414329>.
- Vuori, J. (2021a). Laadullinen sisällönanalyysi. Teoksessa Jaana Vuori (toim.) Laadullisen tutkimuksen verkkokäsikirja. Tampere: Yhteiskuntatieteellinen tietoaarkisto. [Laadullinen sisällönanalyysi - Tietoaarkisto](#). [Viitattu 25.11.2024]
- Vuori, J. (2021b). Tapaustutkimus. Teoksessa Jaana Vuori (toim.) Laadullisen tutkimuksen verkkokäsikirja. Tampere: Yhteiskuntatieteellinen tietoaarkisto. <https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/tutkimusasetelma/tapaustutkimus/>. [Viitattu 4.9.2024]
- Walsh, M. J., McAvoy, J., & Sammon, D. (2022). Grounding data governance motivations: a review of the literature. *Journal of Decision Systems*, 31(sup1), pp. 282–298. <https://doi-org.libproxy.tuni.fi/10.1080/12460125.2022.2073637>.
- Weber, K., Otto, B., & Osterle, H. (2009). One size does not fit all — A contingency approach to data governance. *ACM Journal of Data and Information Quality*, vol.1(1), pp. 1–27. doi: 10.1145/1515693.1515696.
- Weibl, J., & Hess, T. (2020). Turning Data into Value - Exploring the Role of Synergy in Leveraging Value among Data. *Information Systems Management*, vol. 37(3), pp. 227–239. <https://doi.org/10.1080/10580530.2020.1696585>
- Zhang, Q., Sun, X., & Zhang, M. (2022). Data Matters: A Strategic Action Framework for Data Governance. *Information & Management*, vol. 59(4). <https://doi.org/10.1016/j.im.2022.103642>.

LIITE A: KYSELYRUNKO

Tämän kyselyn tarkoituksena on kartoittaa Fingridin datanhallinnan nykytilaa ja kehityskohteita. Kysely toimii pohjatietona tulevalle haastattelulle. Vastaamalla kyselyyn ja osallistumalla haastatteluun autat selvittämään, miten voimme Fingridillä hallita ja hyödyntää dataa paremmin ja mitä kehitystarpeita eri tahot pitävät tärkeinä. Kaikki näkemykset ovat arvokkaita, eikä vääriä vastauksia ole. Vastaa siis rohkeasti ja avoimesti omasta näkökulmastasi.

Kyselyn vastauksia käytetään Fingridille kirjoitettavan diplomityön aineistona. Vastauksesi käsitellään luottamuksellisesti, eikä yksittäisiä vastaajia voida tunnistaa lopullisessa raportoinnissa. Vastaajan nimi kuitenkin tallennetaan, jotta kyselyn vastaukset voidaan yhdistää oikeaan haastateltavaan ennen haastattelua. Vastaukset hävitetään diplomityön valmistumisen jälkeen tai viimeistään kuuden kuukauden päästä haastattelusta. Kyselyyn vastaaminen vie noin 15 minuuttia. Jos datanhallinnan käsite ei tunnu tutulta, alta löydät siitä lyhyen kuvauksen. Kiitos osallistumisestasi!

Datanhallinnan (data governance) tarkoitus on varmistaa, että yrityksen tärkeää dataa hallitaan järjestelmällisesti ja parhaiden käytäntöjen mukaisesti. Se sisältää strategioiden, politiikkojen sekä organisoitumisrakenteiden ja vastuiden määrittelyn. Datanhallinta esimerkiksi määrittää, miten data- ja analytiikkakehitys on organisoitu Fingridillä ja kenen vastuulla se on. Data- ja analytiikkakehitys taas kattaa käytännön työnohjauksen ja kehitystyön, kuten raporttien määrittelyn ja toteutuksen. Viime vuosina Fingridillä datanhallintaa on kehitetty liiketoimintalähtoisemmäksi. Organisaatio on jaettu tietoaalueisiin, jotka edustavat liiketoiminnan osa-alueita tai prosesseja. Samalla keskitetystä raportti- ja datakehityksestä on siirrytty hajautettuun malliin, jossa vastuu kehityksestä on jaettu liiketoiminnan osa-alueille. Tavoitteena on ollut tuoda data- ja analytiikkaresurssit lähemmäs liiketoimintaa. Organisaatiossa on myös otettu käyttöön uusia datanhallinnan rooleja, kuten tietoaalueen omistaja, sekä muodostettu datanhallintaryhmä, joka vastaa datahallinnan kehityksestä koko organisaation tasolla.

Olen

a) ylintä johtoa b) päällikkö c) kehitystiimin vetäjä d) asiantuntija

1. Miten kauan olet ollut tekemisissä datan ja analytiikan kehittämisen tai hyödyntämisen kanssa?

a) alle 1 b) 1–5 c) 5–10 d) yli 10

2. Miten kauan olet ollut tekemisissä datanhallinnan kehittämisen kanssa?

a) alle 1 b) 1–5 c) 5–10 d) yli 10

Datanhallinnan merkitys ja tavoitteet:

3. Arvioi seuraavia väittämiä (täysin eri mieltä / osittain eri mieltä / neutraali / osittain samaa mieltä / täysin samaa mieltä / en osaa sanoa).
 - a. Pidän datan ja analytiikan hyödyntämistä erittäin tärkeänä Fingridin liiketoimintatavoitteiden saavuttamisen kannalta seuraavan vuoden aikana.
 - b. Pidän datan ja analytiikan hyödyntämistä erittäin tärkeänä Fingridin liiketoimintatavoitteiden saavuttamisen kannalta seuraavan 5-10 vuoden aikana.
 - c. Pidän datanhallinnan toimintamallin kehittämistä erittäin tärkeänä Fingridille.
4. Alla on lueteltu asioita, joita datanhallinta voi mahdollistaa. Aseta ne tärkeysjärjestykseen (1=tärkein).
 - a) Strategian toimeenpano
 - b) Riskien vähentäminen, esim. toiminnan säännöstenmukaisuuden varmistamisen kautta
 - c) Toiminnan / prosessien optimointi ja ennustekyvyn kehittäminen
 - d) Toiminnan / prosessien seuraaminen ja raportointi
5. Arvioi seuraavaa väittämää (täysin eri mieltä / osittain eri mieltä / neutraali / osittain samaa mieltä / täysin samaa mieltä / en osaa sanoa).
 - a. Mielestäni Fingridin datanhallinta on linjassa yrityksen strategian kanssa.

Datanhallintamallin selkeys ja soveltuvuus

6. Arvioi seuraavia väittämiä (täysin eri mieltä / osittain eri mieltä / neutraali / osittain samaa mieltä / täysin samaa mieltä / en osaa sanoa).
 - a. Fingridin data- ja analytiikkakehitys on liiketoimintalähtöistä
 - b. Oma työni vaikuttaa Fingridin kykyyn hyödyntää dataa ja analytiikkaa
 - c. Roolini on selkeä datan ja analytiikan hyödyntämiseen liittyen
 - d. Fingridin datanhallinta on liiketoimintalähtöistä
 - e. Oma työni vaikuttaa Fingridin datanhallintaan
 - f. Roolini on selkeä datanhallintaan liittyen

7. Kumpaa pidät tärkeämpänä Fingridin liiketoiminnalle datan hyödyntämiseen ja hallintaan liittyen? (Vaihtoehto A selkeästi tärkeämpi / Vaihtoehto A hieman tärkeämpi / Vaihtoehto B hieman tärkeämpi / Vaihtoehto B selkeästi tärkeämpi)
- A) datan turvaaminen vai B) datan tehokas hyödyntäminen
 - A) kontrolli – B) joustavuus
 - A) säännöstenmukaisuus – B) innovatiivisuus
8. Arvioi, miten seuraavat asiat toteutuvat omassa toimintaympäristössäsi. (täysin eri mieltä / osittain eri mieltä / neutraali / osittain samaa mieltä / täysin samaa mieltä / en osaa sanoa).
- Data löytyy luotettavasti
 - Datan hyödyntäminen on helppoa
 - Dataratkaisuiden tekeminen on nopeaa

Datanhallinnan toteutus

9. Arvioi seuraavia väittämiä (täysin eri mieltä / osittain eri mieltä / neutraali / osittain samaa mieltä / täysin samaa mieltä / en osaa sanoa).
- Minulla on selkeä tilannekuva työni kannalta oleellisesta kehityksestä datan ja analytiikan hyödyntämiseen liittyen.
 - Tieto dataan ja sen hallintaan liittyvien käytänteiden ja menettelytapojen muutoksista saavuttaa minut.
 - Koen, että datanhallintaan osallistuvilla henkilöillä on tarpeeksi aikaa tehtäviensä toteuttamiseen.
 - Koen, että organisaatiossa on tarpeeksi ymmärrystä ja osaamista ja datanhallintaan liittyen.
 - Mielestäni eri tiimit tekevät tarpeeksi yhteistyötä datan hallinnan ja käytön suhteen.

Datanhallinnan vaikutukset

10. Arvioi seuraavia väittämiä (täysin eri mieltä / osittain eri mieltä / neutraali / osittain samaa mieltä / täysin samaa mieltä / en osaa sanoa).
- Olen kokenut positiivisia vaikutuksia data- ja raporttikehityksen vastuun jakamisesta liiketoimintayksiköille.
 - Mielestäni datanhallinta luo arvoa Fingridille. (1–5)

c. Pidän selkeänä, mitä hyötyä datanhallinta tuottaa Fingridille. (1–5)

LIITE B: HAASTATTELURUNKO

1. Oliko sinulle selkeä kyselyssä käytetty jako datanhallinnan ja käytännön data- ja analytiikkakehityksen välillä?
2. Vastasit kyselyyn ennen haastattelua, heräsikö siitä jotain kysymyksiä tai ajatuksia, joita haluaisit tuoda esiin heti alkuun?
3. Miten kuvaisit rooliasi datanhallintaan liittyen?
4. Mitä datan ja analytiikan hyödyntämiseen liittyviä vastuita tunnistat itselläsi?
5. Forms, kysymys 7, "Fingridin data- ja analytiikkakehitys / datanhallinta on liiketoimintalähtöistä.". Voitko antaa esimerkkejä, miksi koet näin?
6. Miten koet, että Fingridillä pystytään tällä hetkellä luomaan arvoa datasta?
7. Oletko tunnistanut esteitä tai hidasteita datan hyödyntämiselle?
8. Millainen datanhallinnan tila Fingridillä on mielestäsi tällä hetkellä? Tunnistatko siinä haasteita tai kehityskohteita?
 - a. Forms, 10. Jos kyselyn vastauksissa nousee joku toteuttamisen osa-alue haasteeksi, pyydetään vastaajaa perustelemaan, miksi kokee näin.
9. Forms, kysymys 5, "Alla on lueteltu asioita, joita datanhallinta voi mahdollistaa. Aseta ne tärkeysjärjestykseen (1=tärkein)". Laitoit datanhallinnan mahdollistamat asiat seuraavaan tärkeysjärjestykseen. Pystytkö perustelemaan, miksi valitsit näin? Toteuttaako mielestäsi Fingridin datanhallinta näitä tehtäviä ja kuinka hyvin?
10. Forms, kysymys 6, "Mielestäni Fingridin datanhallinta on linjassa yrityksen strategian kanssa." Tarkistetaan, mitä vastaaja on vastannut ja kysytään perusteluja, jos ääripää.
11. Forms, kysymys 8, "Kumpaa pidät tärkeämpänä Fingridin liiketoiminnalle datan hyödyntämiseen ja hallintaan liittyen". Arvioit tärkeämmäksi seuraavat tekijät. Pystytkö avaamaan, miksi päädyit tällaisiin vastauksiin ja toteutuvatko tekijät mielestäsi riittävästi Fingridin datan hyödyntämisessä ja hallinnassa?
12. Forms, kysymys 9, "Arvioi, miten seuraavat asiat toteutuvat omassa toimintaympäristössäsi.". Onko arvioimasi asiat riittävällä tasolla omassa ympäristössäsi? Mitä tarvittaisiin lisää ja miksi?
13. Forms, kysymys 11, "Olen kokenut positiivisia vaikutuksia data- ja raporttikehityksen vastuun jakamisesta liiketoimintayksiköille.". Voitko avata, miksi vastasit

kysymykseen näin. Mitä vaikutuksia olet huomannut? Pystytkö antamaan esimerkkejä omasta työstäsi tai organisaatiotasolta?

14. Forms, kysymys 11, "Mielestäni datanhallinta luo arvoa Fingridille". Jos arvo tunnistettu: Mitä arvoa sinun mielestäsi datanhallinta tuottaa Fingridille? Voitko antaa esimerkkejä tilanteista, joissa olet huomannut datanhallinnan tuottavan hyötyjä?
15. Forms, kysymys 1, "Pidän selkeänä, mitä yötystä datanhallinta tuottaa Fingridille". Jos datanhallinnan hyödyt ovat selkeitä, mihin se perustuu (oma fiilis / mittarit / viestintä..), jos ei, mikä auttaisi tuomaan hyödyt näkyväksi?
16. Mitä toivot, että datanhallinta mahdollistaisi tai parantaisi Fingridillä tulevaisuudessa?
17. Onko sinulla vielä jotain lisättävää tai haluaisitko kommentoida jotain muuta datanhallintaan liittyen?